Redacted Version of Jonathan Hochman 4.15.2022 Expert Report and all Appendices (Plaintiffs)

[Document sought to be sealed]

Contains Material Designated "Highly Confidential – Attorneys' Eyes Only" by Plaintiffs

IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA

CHASOM BROWN, WILLIAM BYATT, JEREMY DAVIS, CHRISTOPHER CASTILLO, and MONIQUE TRUJILLO, individually and on behalf of all similarly situated,	§	
Plaintiffs,	§	**Case No. 4:20-cv-03664-YGR-SVK**
v.	§	
GOOGLE LLC,	§	
Defendant.	§	

EXPERT REPORT OF JONATHAN E. HOCHMAN

April 15, 2022

TABLE OF CONTENTS

I. EXECUTIVE SUMMARY OF OPINIONS

1. Pursuant to the Court's Standing Order, this section includes an executive summary of each opinion to be proffered.

2. Opinion 1: As described in Section VIII.A, it is my opinion that, throughout the class period, Google, by way of various tracking beacons, intercepted private browsing communications between users and non-Google websites while those communications were in transit.

3. Opinion 2: As described in Section VIII.A, it is my opinion that a major function of the tracking beacons is to collect highly personal information about users' browsing activities, including users who choose a private browsing mode, such as the contents of their communications with non-Google websites in the form of detailed URL requests, webpage and video interactions, and more.

4. Opinion 3: As described in Section VIII.A, it is my opinion that, throughout the class period, the Google tracking beacons which cause private browsing communications to be intercepted neither facilitate nor are incidental to the communications between users and non-Google websites. The Google tracking beacons functioned (and continue to function) to listen in and collect information from these private browsing communications.

5. Opinion 4: As described in Section VIII.A, it is my opinion that Google could have at any point before or during the class period, redesigned Chrome Incognito to either stop or limit Google's collection of private browsing information from the private communications between users and non-Google websites.

6. Opinion 5: As described in Section VIII.B., it is my opinion that Google, throughout the class period, intercepted private browsing communications without notifying users or providing a choice at the time of collection. Google could have provided such notification but did not.

7. Opinion 6: As described in Section VIII.C., it is my opinion that Google, throughout the class period, intercepted private browsing communications without notifying websites or providing a choice at the time of collection. Google could have done that but did not.

8. Opinion 7: As described in Section VIII.D., it is my opinion that Google, throughout the class period, has stored the private browsing information it has collected in many Google data sources, sometimes permanently.

9. Opinion 8: As described in Section VIII.D, it is my opinion that Google, within the Special Master process, did not identify all logs or data sources that contain private browsing information, omitting key logs.

10. Opinion 9: As described in Section VIII.D, it is my opinion that Google, throughout the class period, has not provided an option for users to review or request deletion of their private browsing information. Google only provided such controls for activities when users are signed into Google, while giving private browsing users no controls when they are not signed into Google.

11. Opinion 10: As described in Section VIII.E, it is my opinion that Google, throughout the class period, created detailed profiles tied to various Google identifiers (that remain undisclosed to users) based on the private browsing information it collected.

12. Opinion 11: As described in Section VIII.E, it is my opinion that Google, throughout the class period, collected valuable data (including private browsing information collected from users' visits to non-Google websites while signed out of any Google account) to feed and develop Google's machine learning algorithms, which Google in turn uses to generate advertising revenues.

13. Opinion 12: As described in Section VIII.E, it is my opinion that Google's tracking beacons have, throughout the class period, enabled Google to serve advertisements to users visiting non-Google websites within their private browsing sessions while not signed into any Google account.

14. Opinion 13: As described in Section VIII.E, it is my opinion that Google, throughout the class period, used private browsing information to serve personalized advertisements to private browsing users, including on non-Google websites and while not signed into any Google account, using their private browsing information.

15. Opinion 14: As described in Section VIII.E, it is my opinion that Google, throughout the class period, used private browsing information to measure and model conversions.

16. Opinion 15: As described in Section VIII.E, it is my opinion that Google before and during the class period attempted to circumvent efforts by other companies to block Google tracking beacons (including with Firefox private browsing's Tracking Protection and Enhanced Tracking Protection) and to block tracking cookies (including with Apple's Intelligent Tracking Prevention).

17. Opinion 16: As described in Section VIII.E, it is my opinion that Google, throughout the class period, used private browsing information collected through Google Analytics tracking beacons for purposes of delivering advertisements, including personalized advertisements and remarketing.

18. Opinion 17: As described in Section VIII.E, it is my opinion that Google, throughout the class period, included private browsing information in the data sets used to enhance Google Search revenues and to develop and improve other Google products and algorithms.

19. Opinion 18: As described in Section VIII.F, it is my opinion that, throughout the class period, information tied to a user's Google account could be linked to the same individual's private browsing information stored within Google logs and data sources.

20. Opinion 19: As described in Section VIII.F, it is my opinion that, throughout the class period, information tied to a user's account with non-Google websites could be linked to the same individual's private browsing information stored within Google logs and data sources.

21. Opinion 20: As described in Section VIII.F, it is my opinion that Google's deletion of certain data (including data Google deleted after this lawsuit was filed) hinders the process of linking (or joining) a user's private browsing information to that user's Google account, and in some cases may make it impossible to do so where it would have been possible but for Google's deletion of the data.

22. Opinion 21: As described in Section VIII.G, it is my opinion that, throughout the class period, Google's systems have provided Google's California-based employees with access to private browsing information that Google collected from users nationwide, with Google employees routinely accessing that information in California.

23. Opinion 22: As described in Section VIII.H, it is my opinion that there are several ways to use Google's records to identify class members. For example, for Class I, the Chrome class, Google's records can be used to (1) identify Incognito traffic, and (2) link that Incognito traffic to users' Google accounts or to users' accounts with non-Google websites. As another example, for both Classes, if users were to come forward and self-identify, Google's records can be used to verify that the individual is a Google account holder who visited a non-Google website that contains Google tracking beacons.

24. Opinion 23: As described in Section VIII.H, it is my opinion that Google's "maybe_chrome_incognito" bit reliably identifies Incognito traffic.

25. Opinion 24: As described in Section VIII.H, it is my opinion that Google's deletion of certain data throughout the class period will hinder the class identification methods I propose, and in some cases may make it impossible to identify certain class members.

26. Opinion 25: As described in Section VIII.H, it is my opinion that Google obstructed the parties' discovery efforts within the Special Master process, including by failing to identify logs that contain private browsing information, producing incomplete log schemas that omitted key fields, not properly formatting search queries and parameters, and delaying productions.

27. Opinion 26: As described in Section VIII.I, it is my opinion that Google, throughout the class period, uniformly attempted to intercept all private browsing communications with non-Google websites that have a Google tracking beacon—regardless of which private browsing mode the user employed.

28. Opinion 27: As described in Section VIII.I, it is my opinion that Google's tracking beacons were so ubiquitous throughout the class period that there is a near certainty that almost every person using the private browsing modes at issue (in Chrome, Safari, and Edge/IE) during the class period had their private browsing information intercepted by Google, including while visiting non-Google websites without being signed into any Google account.

29. Opinion 28: As described in Section VIII.I, it is my opinion that Google does not offer users any control to escape Google's tracking beacons, which are almost impossible to avoid for Internet users in the United States.

30. Opinion 29: As described in Section VIII.J, it is my opinion that, throughout the class period, Google's Chrome Incognito mode functioned in ways that differed from how Google represented to users that Chrome Incognito functions.

31. Opinion 30: As described in Section VIII.K, it is my opinion that Google, before or at any point during the class period, could have changed Chrome Incognito mode to address what Google referred to as user misconceptions, but Google did not.

32. Opinion 31: As described in Section VIII.L, it is my opinion that Google could delete from its systems records of Chrome Incognito browsing communications. Google could also delete from its systems records of users' signed-out browsing communications.

II. INTRODUCTION

33. In the United States, Google has for many years been the leading online advertising platform. To support its online advertising business, Google uses tracking beacons on non-Google websites. Those Google tracking beacons enable Google to collect information when people browse the Internet, including information sufficient to identify users and their devices, with high probability, and to provide Google with a copy of the content that the user has requested from non-Google websites. Google's tracking beacons enable Google to collect information and monitor each user's interaction with a non-Google website in real time.

34. Google's tracking beacons are ubiquitous, present on [redacted] websites, including many of the top non-Google websites and many high-traffic sensitive websites. Examples include: CNN, The New York Times, Tinder, Grindr, Planned Parenthood, Yelp, Zillow, Reddit, LinkedIn, Alcoholics Anonymous, eBay, National Suicide Prevention Hotline, Lyft, Uber, Accuweather, the National Rifle Association, Kaiser Permanente, Pornhub, Weedmaps, and PBS. I provide a list over 80 such websites in Appendix C.

35. Google offers a product called Chrome, which is a browser that people can use to browse online and visit websites. A feature of Google's Chrome browser is the "private browsing" mode that Google has named "Incognito" mode, which Google has offered continuously since 2008.

36. Throughout the class period, one way in which Google has benefited from its Chrome browser has been by protecting the functionality of Google's tracking beacons, allowing Google to collect the data that is the lifeblood of Google's business.

37. Information collected with Google's tracking beacons is beamed directly to Google while web pages are loaded from a third-party website (i.e., a non-Google website) by the user's browser. Importantly, throughout the class period, Google tracking beacons sent data to Google even when a user is in one of the private browsing modes at issue in this lawsuit (in Chrome, Safari, Edge, and Internet Explorer).

38. Based on my analysis of Google's collection, storage, and use of private browsing information during the class period, as compared to Google's disclosures, it is not surprising to me that Google's employees internally expressed concerns. While Google promises privacy and control, throughout the class period, Google in fact tracked and profiled users while they were browsing privately on non-Google websites.

39. For example, Google employees internally wrote that Chrome Incognito was "effectively a lie" (GOOG-BRWN-00630517) and "has always been a misleading name" (GOOG-BRWN-00475063 at -065) with "people operating under a false sense of privacy" (GOOG-BRWN-00457255). Google employees described Incognito as "broken" (GOOG-BRWN-00441285), a "confusing mess" (GOOG-CABR-03827263), and "not truly private" (GOOG-BRWN-00406065). Google employees recognized, "If we don't take [Incognito] on, it will fester and perhaps metastasize, and we will feel like we were derelict in our duty" (GOOG-BRWN-00630517). One employee even said, "I am going to get back on my old shit of yelling that we need to stop calling it Incognito and stop using a Spy Guy icon" (GOOG-BRWN-00475063 at -065).

40. Google employees said, "the blame for people's misconceptions about Incognito mode is due to that name and branding" (GOOG-CABR-03923580) and noted that the "misconceptions flow directly from the framing of the feature as 'Incognito' (or, for other browsers, 'Private')" (GOOG-CABR-04261880).

41. In contrast to these internal discussions, Google publicly asserts that users can control their privacy with Chrome Incognito and the other private browsing modes.

42. I understand that Plaintiffs have alleged that their contract with Google includes the Google Privacy Policy and the Google "Search & browse privately" webpage (GOOG-CABR-00060364)

43. As examples of what Google has told the public, Google's Privacy Policy states:

- "[A]cross our services, you can adjust your privacy settings to control what we collect and how your information is used."
- "You can use our services in a variety of ways to manage your privacy. For example, . . . [y]ou can [] choose to browse the web in a private mode, like Chrome Incognito mode."
- "Our services include . . . [p]roducts that are integrated into third-party apps and sites, like ads, analytics. . ." (https://policies.google.com/privacy?hl=en-US, last accessed April 11, 2022).

44. Similarly, Google's Search & browse privately webpage in the Google Help Center tells the public that "You're in control of what information you share with Google when you search. To browse the web privately, you can use private browsing" (https://support.google.com/websearch/answer/4540094?hl=en, last accessed April 11, 2022). But as described in Section A of this report, Google's tracking and advertising code embedded within non-Google websites collect users' browsing information regardless of whether they choose to use a private browsing mode.

45. I also understand that Plaintiffs have alleged that their contract with Google includes the Chrome Privacy Notice. Within the Google Chrome Privacy Notice, Google states that "Chrome won't store . . . Basic browsing history information like URLs, cached page text, or IP addresses of pages linked from the websites you visit" (https://www.google.com/chrome/privacy/, last accessed April 11, 2022). In fact, as I describe in Section A and elsewhere in this report, Google tracking beacons use the Chrome browser to copy and send users' browsing history to Google servers. This process involves Chrome, for a time, storing that information. And this happens even when a user chooses to browse in Incognito. Google continues to collect and store users' private browsing information even when the users browse in Incognito mode.

46. Google documents state that "Users want to be able to browse the web without feeling as though they are being tracked or having to sacrifice their privacy to do so", and that "Users want access to useful features and experiences, and they want to feel as though the privacy/value trade off is intuitively a fair one (users want to trust that Chrome is doing the right thing by default)" (GOOG-CABR-05248101 at -104). Meanwhile, aside from a fresh set of cookies, Google tracks private browsing users in the same way, using tracking beacons, cookies and parameters embedded in URLs (also called link decorations), as when users are in non-private browsing mode.

47. Google detects when users are in Chrome Incognito mode, and Google could limit its collection, storage, and use of private browsing information, yet Google has programmed its tracking beacons to systematically collect private browsing information for Google and has failed to provide users or non-Google websites an option that would permit Google to only collect non-private browsing information.

48. Throughout the class period, Google could have programmed Chrome to disable Google's tracking beacons when a user enters Incognito private browsing mode or taken other steps to

prevent Google's storage and use of private browsing information collected from other browsers, but Google did not do that and still has not done this. Even after this lawsuit was filed, Google continued to collect, store, and use private browsing information.

49. Google, other search engines, and website publishers have together adopted the Robot Exclusion Standard that prevents data from being harvested from websites by search engine robots, such as Google's Googlebot, when a publisher opts out.[1] Google could have worked with other browser vendors to create an analogous standard protocol to prevent data from being harvested from users' browsers (including non-Chrome browsers) by Google's tracking beacons when a user chooses a private browsing mode. The Robot Exclusion Protocol was established in 1994. Twenty-eight years later, there is still no analogous standard for protecting users from unwanted data harvesting, with Google systematically harvesting private browsing information.

50. A specific example of Google failing to work with other browser vendors to establish privacy standards is the universal Do Not Track ("DNT") signal, which was first proposed in 2009.[2] DNT is a proposed technical standard, analogous to the Robot Exclusion Protocol, that would allow users to signal that they do not want their data harvested. However, despite Google's leading position in the Internet advertising industry and Google's position as the leading browser vendor, Google never adopted anything like the DNT signal for its Google tracking beacons with private browsing. It has been over thirteen years since Google launched Chrome with Incognito mode, and Google has had thirteen years to address the need for a technical means to provide users with a way to control Google's collection, storage, and use of private browsing information, and

[1] https://developers.google.com/search/blog/2019/07/rep-id (Last accessed on April 11, 2022).

[2] https://www.eff.org/issues/do-not-track, (Last accessed on April 11, 2022).

to provide users with privacy when they chose to use a private browsing mode. As the leader in Internet advertising, Google has failed to do so.

51. Throughout the class period, voluminous amounts of private browsing user data transmitted to Google by Google's tracking beacons was received and saved by Google.

52. Throughout the class period, Google used this private browsing information for Google's benefit, including to increase Google's profits. For example, information about the content viewed by users (including while browsing privately) allows Google to target ads more precisely in support of higher advertising rates. By tracking user purchases with "conversion tracking" (including while browsing privately) Google can substantiate the value of advertisers' expenditures. Conversion tracking data also enables Google's automated bidding algorithms that help advertisers bid more efficiently (and spend more money with Google). In addition, Google's advertising revenue for certain kinds of ads depends upon collecting documentation of user behavior, such as watching some percentage of a video ad.

III. STATEMENT OF LIMITATIONS REGARDING THIS REPORT

53. I prepared this report for purposes of this case only. It may not be used for any other purpose. This report contains and refers to information designated as "CONFIDENTIAL" and "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY" under a Stipulated Protective Order.

IV. ENGAGEMENT

54. Counsel for the Plaintiffs in this action ("Counsel") retained me to develop and render opinions concerning the technology and practices at issue in this litigation with respect to several products (the "Google Products" as described in Appendix A) developed and distributed by defendant Google, LLC ("Google"), including based on my analysis and review of documents, testimony, data produced through the Special Master process, Google Product tests and publicly

available code. The Google Products include those utilizing Google tracking code (e.g., Google Analytics and conversion tracking code) and Google advertising code (e.g., Google Ad Manager and Google AdSense advertising code).

55. I am charging an hourly rate of $800/hour and my associate, Julie Ann Burns, is charging an hourly rate of $200/hour. Our compensation does not depend upon the outcome of the case. In the event of any recovery in this case, I understand that Ms. Burns and I will be excluded from any disbursement of funds.

56. While I have used many browsers in private browsing mode, both before and throughout the class period while not signed into any Google account and to visit non-Google websites, and while I am a Google Analytics and Google Ads customer and a Google Account holder, before this engagement, I was not aware of Google's collection, storage, and use of private browsing data. I only became aware of that through discovery in this case.

57. I have had full access to every document produced in this case, using a document review platform provided by International Litigation Services ("ILS"), and also to all of the deposition transcripts, written discovery responses, and data and information produced through the Special Master process. I was freely allowed to conduct searches and view any document for the purpose of preparing this report, and I spent many hours independently searching for and reviewing documents produced in this case and reviewing the deposition transcripts, written discovery responses, and data and information produced through the Special Master process.

58. To the extent any other documents or data are produced or there is any further testimony, I reserve the right to supplement this report.

V. EXPERTISE

59. I have a Bachelor of Science (BSc.) degree awarded *summa cum laude* and Master of Science (MSc.) degree in computer science, both from Yale University. I earned my degrees in 1990, one of four students in my class of 1500 to receive simultaneous degrees. My master's thesis, entitled *How to Create a Failure Tolerant Distributed System*, focused on distributed transaction processing and the distributed consensus problem. My coursework at Yale included theory of computation, distributed systems, artificial intelligence, systems programming, and operating systems. My thesis advisor was Dr. Michael J. Fischer, a renowned computer science professor known for his work in distributed computing. I was encouraged to attend Yale by, and then studied with, Dr. Alan Perlis, the first recipient of the Turing Award, the "Nobel Prize of computing."

60. Over the last 30 years I have maintained regular correspondence and meetings with my advisor, Dr. Fischer, which helps me to stay informed about the latest academic work in computer science. During the fall semester of 2020 I took his course CPSC 467: Cryptography and Security at Yale, and during the spring semester of 2021 I took his course CPSC 457: Sensitive Information in a Connected World. The latter course placed a heavy emphasis on the ethical issues related to the handling of other people's data.

61. Dr. Fischer, Dr. Daniel Boffa, and I recently authored an academic paper, *Privacy-Preserving Data Sharing for Medical Research,*[3] that I presented at the 23rd International Symposium on Stabilization, Safety, and Security of Distributed Systems in November 2021. I also presented our work at the Yale Applied Cryptography Laboratory on October 22, 2021. This paper relates to the ethical handling of other people's sensitive data and computer security in general.

[3] https://cpsc.yale.edu/sites/default/files/files/TR1558.pdf

62. In August 2022, I will become a Doctoral Candidate at Yale University in the Computer Science Department. My academic work will be funded by Yale University for at least the first year of my studies.

63. I serve as an expert consultant and expert witness in fields relevant to my expertise and experience, such as software development, technology entrepreneurship, and Internet advertising, marketing, e-commerce, and security. A complete copy of my *curriculum vitae* is attached as **Exhibit A**. It includes a list of prior cases in which I have testified as an expert at trial or deposition. As of the date of this report, I have testified in 11 trials and at 39 depositions. I have been qualified as an expert in two United States District Courts, as well as state courts in California and Connecticut, four arbitrations, and one court in Israel.

64. I am the founder of Hochman Consultants, an Internet marketing agency, established in 2004. Hochman Consultants provides clients with marketing strategy, website development, ecommerce, digital advertising, and Internet security services. Hochman Consultants also provides advice on search engine optimization and pay-per-click advertising, making websites more user friendly, increasing traffic, and increasing the value of each visitor. I am a member of the Association for Computing Machinery, the world's largest computing society, whose goal is to advance computing as a science and a profession.

65. I also am a co-founder of the UNS Project, an Internet security venture to build network infrastructure services for (a) user authentication, (b) privacy preserving record linking (PPRL), and (c) Sybil authentication, the detection of fake accounts. I recently gave a presentation about the UNS Project at the Yale Innovation Summit in May 2019.

66. I was a co-founder of CodeGuard, a startup information security company. In July 2018 CodeGuard was acquired by Sectigo, an information security company controlled by Francisco

Partners, for an undisclosed price. CodeGuard had more than 250,000 paying customers as of 2018. The CodeGuard product has since been integrated into the Sectigo Web Security Platform which is sold worldwide.

67. From 1999–2004, I served as a Director of Barcoding Inc., a systems integrator, which develops and provides software, mobile applications, mobile computing, automatic data collection, and wireless networking. While I was a Director, Barcoding Inc. was named to the Inc. 500, recognized by Forbes Magazine as one of ten privately held companies to watch, and ranked third on the Maryland Fast 50. During my work at Barcoding Inc. I was responsible for the company's e-commerce platforms.

68. I presently hold Google Ads search and display certifications. To obtain these certifications I studied Google's training materials and passed Google's exam. I have held Google Ads certifications and repeated this process when required to renew my certifications.

69. I work for clients in diverse industries and government roles. My clients have included major corporations and organizations, such as Apple (on two occasions), Zillow, Delta Air Lines, U-Haul, the FDIC, and the US Treasury. I have written scores of expert reports for both plaintiffs and defendants alike without preference for one side or the other. Many of these reports are not reflected on my CV because many cases settle without my being called to testify.

VI. PREPARATION

70. I spent hundreds of hours reviewing materials produced in this case, including documents that I selected from a database containing over 900,000 documents as well as deposition transcripts and written discovery responses.

71. I was actively involved in the Special Master process, through which I obtained access to information and data concerning Google's collection, storage, and use of private browsing information during the class period.

72. I also reviewed public materials, such as Google support pages and publicly available source code for the Google Products.

73. I reviewed user interfaces of the Chrome, Firefox, Safari, Edge, and Internet Explorer browsers.

74. I conducted experiments with a web debugging and proxy tool called Fiddler Classic to assess how the Google Products function. I also conducted experiments using browser web developer tools.

75. Through the Special Master process, tests have been conducted on named Plaintiff and consultant Mr. Thompson and Dr. Dai's devices to better understand Google's logging and storage systems. These tests have been conducted by the Plaintiffs as well as by consultants accessing these devices under my direction. Due to limited travel during Covid, we have sometimes remotely accessed these devices using a remote desktop software called TeamViewer.[4]

76. I was assisted in my work by the following consulting experts, who I supervised: Dr. Lillian Dai, Mr. Jay Bhatia, Mr. Vivek Shinde, Mr. Christopher Thompson, Mr. Jake Taylor, Concur IP engineers, and Ms. Julie Ann Burns. I also had conversations with Mr. Michael Lasinski and members of his team, Mr. Steven Weisbrot, and Mr. Bruce Schneier. This report was the culmination of work performed by me or by others working under my direction and direct supervision. All opinions rendered in this report are my own.

VII. BACKGROUND / ASSUMPTIONS

77. For purposes of my analysis and opinions, I assumed the following:

1) The operative complaint is the Third Amended Complaint, which is Docket number 395-2.

2) The relevant period started on June 1, 2016 and is ongoing (the "class period").

[4] https://www.teamviewer.com/en-us/ (Last accessed on April 11, 2022).

3) Plaintiffs assert claims on behalf of US-based Google Account holders.

4) Plaintiffs allege two classes: the Chrome Incognito class, and the non-Chrome private browsing class (which I understand based on discussions with Counsel to be limited to Safari and Edge/Internet Explorer).

VIII. OPINIONS

A. Google Interception: Throughout the class period, Google intentionally intercepted private browsing communications between users and non-Google websites while those communications were in transit

78. Based on my review of documents and testimony as well as experiments with all of the private browsing modes at issue, on both desktop and mobile devices, it is my opinion that Google, throughout the class period, intentionally intercepted private browsing communications between users and non-Google websites[5] while those communications were in transit and collected private browsing information from those communications. Google intercepted private browsing communications while users were not signed into any Google account. Google's interceptions were systematic and consistent throughout the class period and across class members, regardless of the user's type of browser and device.

79. The Google Products throughout the class period have used certain Google tracking code (including Google Analytics[6], conversion tracking[7], and other Google tracking code of a similar

[5] For the purpose of this report, "non-Google websites" (alternatively, "third-party websites") are websites not owned and operated by Google.

[6] In this report, I will use the term "Google Analytics" or "Analytics" to include both the free version of Google Analytics and the premium paid version, Google Analytics 360. Google Analytics 360 offers more tracking features among other premium services (See e.g., GOOG-BRWN-00490767 at -824, -917, -924, -932 to -935). "Google Analytics" also includes the various generations of the Google Analytics product line, including Classic Google Analytics, Universal Analytics, App + Web, and Google Analytics 4 (GA4).

[7] In this report, I will use the term "conversion tracking code" to refer to Google tracking code embedded in non-Google websites for tracking user activities deemed by websites to be a conversion event (e.g., Analytics, Ad Words Conversion Tracking and Floodlight). Since 2017, conversion tracking code have

nature) and advertising code[8] (including Google Ad Manager, Google AdSense, and other Google advertising code of a similar nature) to accomplish these interceptions and collections. In Appendix A, I provide a historical overview of Google's tracking and advertising products and some of their main tracking and advertising code.

80. To assess how these Google Products and their tracking and advertising code function in connection with private browsing modes (e.g., Chrome Incognito, Safari Private Browsing, and Microsoft Internet Explorer/Edge InPrivate), I conducted certain tests and also reviewed many of the "Confidential" and "Highly Confidential" documents produced by Google in this case. I conducted tests using Chrome, Safari, Microsoft Internet Explorer/Edge, and Firefox browsers on desktop and mobile devices and evaluated the publicly available tracking and advertising code associated with the tests. I also reviewed deposition transcripts from Google engineers and Google's written discovery responses regarding Google's tracking beacons. That testimony and Google's discovery responses also inform and support my opinions.

81. My opinions are further detailed in the subsections below.

---

been consolidated under Google's One Google Tag (OGT) initiative (externally referred to as Global Site Tag) as well as under Google Tag Manager (GTM) (GOOG-CABR-05305495, GOOG-CABR-04419756).

[8] In this report, I will use the term "advertising code" to include tracking beacons from Google Ad Manager, Google Ad Manager 360, Google AdSense, as well as the various generations of these products. Prior to 2018, Google's advertising system on non-Google websites (which form Google's display advertising network) included DoubleClick Reservation and Exchange (DRX) (also known as DoubleClick for Publishers (DFP)), Google Display Network (GDN), Google AdSense and Google Ad Exchange (AdX). In 2018, Google rebranded DRX with the sell-side of AdX (i.e., exchange services for websites that sell their advertising space) as Google Ad Manager (GOOG-CABR-04819485).

A.1 Google designed tracking and advertising code to run on non-Google websites for Google's interception

82. As confirmed by my analysis and tests in connection with this case, Google designed its tracking and advertising code (which I refer to as "tracking beacons"[9]) to be loaded from non-Google websites and then run on the user's device, with Google using such code for Google's interception of the user's communication with the non-Google website to copy and collect information from those communications, while those communications are in transit. This was Google's design and functionality for the entire class period, across all class members.

83. Google intercepts private browsing information through tracking beacons designed by Google to intercept and collect information from communications between users and non-Google websites, including while users are in a private browsing mode. A tracking beacon is implemented on non-Google websites via a piece of code, such as one written in JavaScript, that is provided by Google to be embedded on the non-Google websites. Google's tracking beacons are embedded in both its Publisher and its Advertiser websites, and even in websites of third parties who are neither Advertisers nor Publishers of Google's advertisements.[10] On Publisher websites, these Google tracking beacons track users, their browsing activities, and interactions with advertisements. On Advertiser websites, these Google tracking beacons similarly track users, their browsing activities,

[9] The term "tracking beacon" is also used by internal Google documentations. *See e.g.*, "Visitor visits a website and tracking beacon is sent to GA [Google Analytics] collection" (GOOG-CABR-04141705 at -707) and GOOG-CABR-04678169 referring to analytics and ads tracking as a "dual-beacon system."

[10] Publishers are those websites that sell advertisement spaces on their webpages (e.g., New York Times) and Advertisers are those websites that place advertisements on Publisher websites (e.g., Nike places a Nike shoe advertisement on New York Times). Some Publishers are also Advertisers as they may advertise their own products and services on other non-Google websites (e.g., New York Times may place its subscription advertisement on accuweather.com).

and conversion activities[11] (for example, when a user makes a purchase based on a previously viewed or clicked ad). In addition to Publisher and Advertiser websites, Google tracking beacons associated with Google Analytics are also embedded in non-Google websites that are independent of the Google's advertisement ecosystem to track user browsing activities and gain insights on users and their behaviors.

84. As stated in Google's written discovery responses, "websites choosing to use Google Analytics or Ads Manager services install Google code in their website for the purpose of transmitting data to Google so that Google can provide the desired services" (RFA No. 2; *see also* Google's Resp. to RFA 30 ("third party websites choosing to use Google Analytics or Ads Manager services install Google code in their website for the purpose of transmitting data to Google so that Google can provide the desired services"). The same is true with respect to the other Google Products, and this is how these Google Products functioned throughout the class period. And Google has admitted that "Google designed the code provided to third party websites for the purpose of providing these services" (RFA 30). Google tracking beacons throughout the class period had a common functionality in terms of Google's interception and data collection, not limited to the Analytics and Ad Manager codes.

[11] As I will explain in more detail in Section E, a conversion is an action that an advertiser defines to be valuable to their business. For example, an online purchase, a new sign-up, filling out a web form, downloading an app, or even offline events such as a store visit could all be conversion events. Tracking conversions is one of the primary ways Google demonstrates the effectiveness of its advertisement infrastructure to advertisers.

85. Throughout the class period, when a user visited a non-Google website[12] such as New York Times, a GET message would be sent from the user's browser to the non-Google website server. This GET message contains several pieces of information, including:

- Request URL: the URL of the requested webpage of a specific page, image, video or other resources (for example, here is an URL of a specific page on New York Times: https://www.nytimes.com/2022/01/20/business/economy/inflation-questions-users.html)
- IP address: the IP address of the user's device making the request
- User-agent: a text string that identifies the user's device platform (e.g., Windows, Android, etc.), browser (e.g., Chrome, Safari, IE/Edge, etc.) and browser version
- Referer[13]: the URL of the referring page (i.e., URL of the page on which a link is clicked, or the page from which a script or image is referenced)[14]

86. In response to the GET request communication from the user to the non-Google website, the non-Google website server sends back instructions and data to load the webpage in the browser (i.e., fully rendering or displaying the page in the browser). The page load process may take numerous steps depending on the instructions in the webpage, including additional requests sent to the non-Google website server to obtain text, icons, images and other resources (also known as "assets," e.g., JavaScript files, CSS files, image files).

---

[12] A user visits a non-Google website by either entering the website URL in the address bar of a browser, clicking on a link (as a search result or as a link embedded in another website), or clicking on an advertisement, image, video, or other types of linked resources.

[13] In the HTTP protocol specification, the referrer was accidentally mis-spelled "referer" and then came into such widespread use that it was impractical to make a correction. https://en.wikipedia.org/wiki/HTTP_referer#Etymology (Last accessed on April 11, 2022).

[14] See https://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html (Last accessed on April 11, 2022).

87. To observe Google tracking beacons embedded on non-Google websites that send intercepted private communication information to Google, I visited several non-Google websites in Chrome Incognito mode and used the built-in Chrome Developer Tool to observe the types of Google tracking beacons that executed on these websites. Those are tools designed to provide developers with additional information. In my experience, those are not tools that any typical consumer would use or understand.

88. For example, on washingtonpost.com (identified by Google as one of the top 25 domains with the highest website traffic in the U.S. using Google Ad Manager and Google Analytics - Google's Supplemental Response to Interrogatory No. 5), at least the following Google tracking beacons are used: gtm.js, gtag, analytics.js, and gpt.js. Based on the confidential discovery obtained through this litigation, support documents, and my own experience, I understand that these are Google tracking beacons used by Google Tag Manager, Google's global site tag, Google Analytics, and Google Ad Manager.[15]

[15] Google Tag Manager (utilizing the gtm.js script) is a method for deploying analytics and conversion tracking tags (also known as tracking beacons) https://developers.google.com/tag-platform/tag-manager/web and https://support.google.com/tagmanager/answer/7582054?hl=en. GOOG-CABR-04419756 at -757 describes gtag as a unified tagging system for implementing analytics and conversion tracking code, GOOG-BRWN-00149515 at -515 describes analytics.js as a Google Analytics script (also a tracking beacon), and GOOG-CABR-00059481 at -483 describes gpt.js as a Google Ad Manager tagging library.

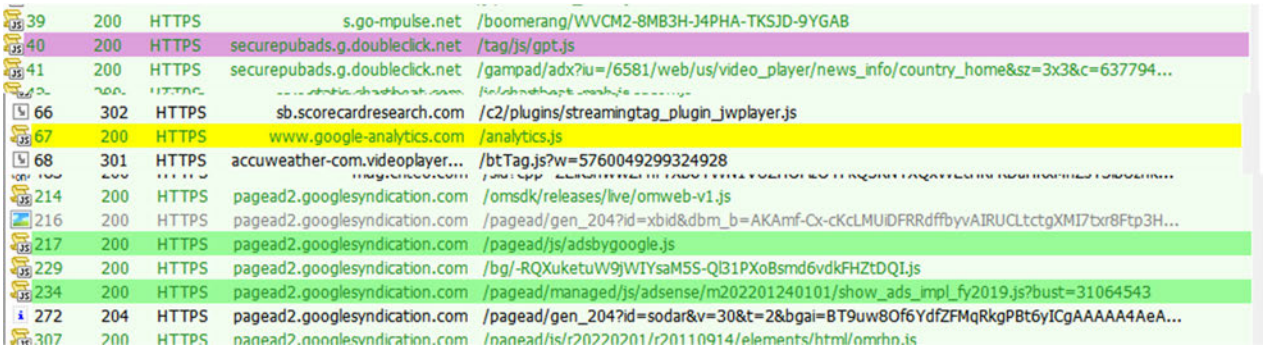
89. As another example, on manning.com, at least the following Google tracking beacons are used: gtm.js, analytics.js, and adsbygoogle.js. These are Google tracking beacons used by Google Tag Manager, Google Analytics and Google AdSense.[16]

90. As another example, on nike.com, at least the following Google tracking beacons are used: gtm.js, gtag, analytics.js, and conversion_async.js. These tracking beacons are used by Google Tag Manager, Google's global site tag, Google Analytics. and for Google's conversion tracking. [17]

[16] GOOG-CABR-04147060 at -065 describes adsbygoogle.js as a Google AdSense tag (yet another tracking beacon).

[17] GOOG-CABR-04077431 at -433 describes conversion_async.js as a conversion script.

91. I also visited websites using various browsers and a HTTP proxy tool called Fiddler Classic to observe Google tracking beacons on Chrome and other browsers in private browsing mode (see Appendix B for details). As an example, on accuweather.com, at least the following Google tracking beacons are used: gpt.js, analytics.js, adsbygoogle.js, show_ads_impl_fy2019.js.[18]



92. Additional examples of various Google tracking beacons are included in Appendix A and Appendix B.

93. In addition, I have examined several websites containing potentially sensitive content and found Google tracking beacons embedded on these non-Google websites as well. Appendix C includes a list of websites that I have examined, and a subset of Google tracking beacons found on

[18] GOOG-CABR-04641725 at -728 describes show_ads_impl_fy2019.js as a Google AdSense script.

the websites. These include websites that provide adult content (pornography), social support (e.g., suicide, addiction, abortion, and family planning), news/media/entertainment, social networking, and dating, government, financial, health, shopping, and transportation services. In most cases, more than one Google tracking beacon was found on the non-Google website.

94. Google's tracking beacons embedded in non-Google websites take up processing, storage, and power/battery resources to run on user devices; thereby increasing user's energy and device costs.[19] These Google tracking beacons also cause webpages to load slower. Out of tens to hundreds of request and response messages sent and received to load a single webpage, it is not uncommon for a significant portion to be intercepted messages going to Google. Depending on the number of requests for assets and the number of tracking beacons on the webpage, a single webpage may take many seconds to load completely.

A.2 Google intercepted (and continues to intercept) private browsing communications between users and non-Google websites, with Google obtaining personal information from those communications

95. I analyzed the functionality of the private browsing modes at issue in the two classes (in Chrome, Safari, and Edge/Internet Explorer), and it is my opinion that Google throughout the class period has intercepted (and continues to intercept) private browsing communications between users and non-Google websites, including with respect to the individuals who would be members in the two alleged classes. When Google tracking beacons are run by one of the browsers at issue on the user's computer, these tracking beacons turn the user's computer into tracking device.

96. In these private browsing modes, when a user's browser requests a webpage from a non-Google website via a GET request, Google's tracking beacons embedded in the web page cause

[19] See also, https://www.mdpi.com/2227-7080/8/2/18/htm (Last accessed on April 11, 2022).

information to be copied from that communication and sent to Google's servers concurrently with the user's private communication with the non-Google website. Webpages typically have numerous assets that are retrieved during the process of rendering the page. In terms of the information copied and collected by Google from these private communications with third parties, this information includes, without limitation, the URL of the specific webpage visited by the user (including the full URL viewed by the user on the non-Google website, which would include folders, subfolders, and precise file requested from the webserver), the user's IP address, and the user agent string of the user's browser, among other information. Google used its tracking beacons to collect this private browsing information throughout the class period.

97. In the case of Chrome Incognito, throughout the class period, Google treated the interception slightly differently by not sending a particular HTTP header called the X-Client-Data header to Google.[20] The lack of the X-Client-Data header indicates to Google that the user could be engaged in Incognito browsing, as recognized by many Google employees.[21] Google also developed algorithms that combine information about the X-Client-Data header with the user agent string to detect whether a user is in Incognito mode.[22] Google's logfiles even include a bit that

[20] *See* Google's Request for Admission No. 12 ("Google further admits that [data generated by Incognito users' visits to websites that use Google's Analytics and Ad Manager services] does not include the X-Client-Data Header."); Google's Response to Interrogatory No. 30 ("If Chrome is used in Incognito mode, Chrome will not send the X-Client Data Header to doubleclick.com or any other domain used by Google Ad Manager or AdSense").

[21] *See* GOOG-CABR-03667366 (noting that the "starting point" to detect use of Incognito mode "will be the (absence of) X-Client-Data header.")

[22] *See* Leung Depo. 47: 13-17 ("My role in this doc is to come up with the heuristic approximation of—I mean, part of that is to come up with the heuristic approximation to—for identifying whether the request may be coming from Chrome incognito); *id.* at 61: 19-21; 62: 2-5 (describing a "proposal" to "compute an enum and if the maybe_Chrome_incognito bit separately than log them into relevant logs" that was "implemented sometime after" the proposal was described in an internal Google document).

indicates whether the user appears to be in Incognito mode.[23] Sections D, F and H of this report and Appendices G to I include additional information.

98. As described in Appendices A and B, when Google tracking beacons execute in private browsing mode in Chrome, Safari, and Edge/Internet Explorer, these Google tracking beacons operate in substantially the same manner in terms of transmitting users' private browsing information to Google server(s) as they do in non-private browsing mode.

99. My analysis and experiments confirmed that Google throughout the class period collected personal information (as detailed below) from intercepted private browsing communications using the Google tracking beacons, including without limitation URLs, IP addresses, user agent, referer, and other information from the user's private browsing communications. Google did this systematically with the private browsing modes for Chrome, Safari, and Edge/Internet Explorer, with common code used for these Google tracking beacons across non-Google websites, notwithstanding whether users chose to browse privately.

100. My opinions are consistent with Google's discovery responses and admissions. In Google's response to Interrogatory No. 30, Google admits that "When a user (in any browser) visits a website that uses Google Ad Manager or AdSense, Google Ad Manager or AdSense may receive: (1) cookies that specific Google domains previously set on the user's browser; (2) the HTTP request sent by the user's browser, including the hostname, browser type, and language, and depending on the browser used, Java support, Flash support, and screen resolution; (3) the URL of the website making the ad request to Google Ad Manager or AdSense, and/or the referrer URL; (4) the IP address assigned to the device on which the browser is running; (5) the request for an ad

[23] Leung Depo 175:8-19 (confirming that browser detection dashboard was likely drawing from maybe_Chrome_incognito bit and other fields added to logs).

to be served on a non-Google website and the ad slot to be filled; (6) event data such as impressions or clicks; and (7) if the user is in Chrome and a mode other than Incognito, the browser's X-Client-Data Header." *See also* Google's Response to RFA No. 52 ("Google admits that if a user visits a webpage that includes at least some of the identified Google AdSense tags, those tags will cause the user's browser to attempt to transmit to Google the URL of the webpage that a user is currently visiting.").

101. In its Answer ¶ 63, "Google admits that when a browser is used to visit a third party publisher website that is using certain Google services, such as Google Analytics or Google Ad Manager, then those Google services may receive information related to that specific browser website visit. Google avers that the information that these Google services receive through this data-flow may include: GET request (referer header and URL), IP address, HTTP headers (which may contain information about the user's browser or device (i.e., "user-agent")), "User-ID" issued by the website publisher, and Google cookies set for relevant domains." (page 8).

102. In response to RFA No. 6: **"**Google admits that, since at least June 1, 2016, depending on (1) the particular website visited, (2) the particular Google service(s) the website uses, (3) the website's specific settings and selections for each of its Google services, and (4) the user settings, browser settings, and other software settings and plug-ins, Google may receive through a service certain data generated by a user's interactions with the service, which data may include: IP address; the web page the user viewed; referral URL; and search queries. Google admits that, to the extent it receives such data, Google receives the data regardless of whether the user is in Incognito mode." And Google's designated 30(b)(6) Google Analytics witness testified that the process of logging information Google receives when the Google Analytics JavaScript executes is "agnostic" to whether the visitor to a webpage is using a private browsing mode. *See* Chung Tr. 20:6-10.

103. In response to Interrogatory No. 1, Google stated that "The fact alone that a user is in private browsing mode does not change the types of data collected when Chrome users visit a third-party website that uses Google Analytics or Ad Manager while not logged in to their Google Accounts . . . Therefore, depending on (1) the particular website visited, (2) the particular Google service(s) the website uses, (3) the website's specific settings and selections for each of its Google services, and (4) the user settings, browser settings, and other software settings and plug-ins, Google may receive through a service certain data generated by a user's interactions with the service, which data may include: IP address; the web page the user viewed; referral URL; and search queries. To the extent it receives such data, Google receives the data regardless of whether the user is in Incognito mode." Moreover, "the foregoing response applies when users visit websites that use Google Analytics and Ad Manager services in non-Chrome browsers' private browsing modes" (*Id.*). Google Chrome engineer Halavati confirmed that Google stores this data it receives when a user is in private browsing mode on Google's servers. Based on my own experience as a Chrome user, and as confirmed by Halavati, Google does not give users the option to delete this data. *See* Halavati Tr. 88:14-89:3. According to Google's internal documents, CEO Sundar Pichai's stated position is that "if you don't want to share your data with Google, Chrome might not be the right browser for you" (GOOG-CABR-00427432 at -447). This position statement does not appear in any public document on the internet indexed by the Google search engine, nor has it been reported in any news source that I can find.

104. In addition, Google employee Brian Rakowski, the so-called "father of Incognito" (Rakowski Tr. 19:19-20:2), testified during his deposition that Google receives an "HTTP request" "in any situation where a request came to a Google server," which include circumstances when the webpage a user requests "embeds content from Google" (e.g., Analytics, Ad Manager, AdSense,

conversion tracking) (Rakowski Tr. 51:3-22). The request received by Google "contains . . . all the standard browser header information, including IP address, user-agent string." *Id.* 81:12-82:2. Mr. Rakowski likewise testified that Google did not design Chrome Incognito mode "to prevent Google from receiving any personally identifiable information." *Id.* 60:8-15. He testified that Chrome Incognito mode "was never about changing logging behavior, period." *Id.* 238:14-18. Mr. Rakowski was also not aware of any controls that allow a user to prevent all server logging by Google while they're browsing in Incognito mode. *Id.* 316:23-317:1; *see also id.* 46:19-25 (explaining that Google receives the IP address and user agent from users that visit Google Analytics or Ad Manager websites within a non-Google private browsing mode).

105. It is my opinion that an IP address, especially when combined with a user-agent string, constitutes personally identifiable information ("PII") because this data can be used to uniquely identify a user with a high probability of success. "Personally identifiable information (PII) is any data that could potentially identify a specific individual. Any information that can be used to distinguish one person from another and can be used to deanonymize previously anonymous data is considered PII,"[24]

106. Google employees also recognized, consistent with my analysis and opinions, that the intercepted and collected information in private browsing mode are sensitive, personal information. Google's internal document acknowledges that "Personal information includes data such as IP address, cookie/device ID, PII, postal address, geo-location data etc." per California Consumer Privacy Act (GOOG-BRWN-00026989 at -991) and "Fingerprinting techniques rely on collecting several unique attributes about a user's device such as operating system, browser

[24] Google Quick Answer, referencing https://www.techtarget.com/searchsecurity/definition/personally-identifiable-information-PII (Last accessed on April 11, 2022).

version, IP address, browser language, fonts installed, screen resolution and more. Combined, those details create a unique profile of a user's device and, by proxy, the user. Users can't see that this data is being collected, nor can they delete it by clearing cookies." (GOOG-BRWN-00026989 at -992). Google engineers wrote "I believe that preventing the data collection is better for the user" and that Google is "making privacy worse by collecting Incognito when the user has clearly told Chrome that it wants to be private. That is the fundamental reason for trying to strengthen Incognito and allow Google services to actually deliver on the promise of being private." (GOOG-CABR-00501220 at -220 and -221). Google employees also recognized that until Google made changes that it never made (including "[r]emove IP address, user agent, and cookie IDs from Incognito data"), Google could not claim that it "does not save personal information from your Incognito session" (GOOG-CABR-00421851 at -851 and -852).

107. Google has also received other personally identifying information and identifiers from private browsing communications with non-Google websites that can be used to identify users. These include various information about users (email, name, phone number, home address, etc. obtained, for example, from forms users fill in on non-Google websites[25]) as well as User IDs ("UIDs") used to track individual users logged into non-Google websites and Publisher Provided IDs ("PPIDs") used to track individual users logged into websites that use Google Ad Manager

[25] GOOG-CABR-05405676 at -691 and -692: Enhanced conversion "works for Web Conversions in Google Ads. EC relies on hashed customer data for matching, attributing and measuring conversions … For the best match results, please consider sending email address in addition to other fields: Email address (recommended), Phone number, Name & Home Address (which is a combination of the below, all must be included – First name, Last name, Street, City, Region, Postal code, Country" and GOOG-BRWN-00549861 at -862: "Enhanced UID enables the use of a User-ID as a SHA256 hashed version of email address, phone number, or mailing address."

360. [26] In addition, Google receives a myriad of fingerprinting information, including details regarding the user's screen resolution, screen color depth, time zone, language, and Internet connection bandwidth/speed.[27] Google collects this information not only from regular browsing but also when users are privately browsing non-Google websites without being signed-in to any Google services. This detailed information obtained by Google from users' private browsing activities allows Google, or somebody who comes to possess Google's data, to know precisely what a person was viewing within a private browsing session.

108. Another important set of information received by Google through these interceptions and throughout the class period is web cookies, which are pieces of "text stored by a user's web browser"[28] containing identifying information and used by Google for tracking and advertising purposes. Google uses the "uniquely generated ID that is stored in a web cookie in browsers to identify a user" (GOOG-BRWN-00844093 at -095). I would give a more precise definition of web cookie than the one Google quoted. A web cookie is a small file sent by a web server or set by a

[26] User IDs include, but are not limited to, PPID, Analytics User ID and Google Ads User ID (also known as CRM-ID). See e.g., GOOG-CABR-04772394 at -394: "(PPID) has existed in the Ad Manager product for a number of years (go/xfpppid from 2012)"; GOOG-CABR-05136994 at -996: "PPID/CRM_ID-based cross device effort is to leverage the login information from 3rd party publishers. 3rd party publishers gain access to user login on their sites and can send requests to doubleclick.net in the form of ad request (A.K.A. PPID) or remarketing request (A.K.A. CRM ID) with the user's login information appended. Subsequently, our network has simultaneous access to both the 3rd party login and the biscotti id and thus can use 3rd party login as the bridge to connect biscotti ids from different devices, apps and browsers."; GOOG-CABR-05673102 at -103: "User ID was called CRM-ID before" and at -104: "When a user visits the advertiser's website, a smart pixel ping is fired to google, user id, account_id and biscotti cookies are sent to EventFE [Event Front End]"; GOOG-CABR-04691939 at -939: "If clients can identify the user (via user ID), they'll be able to use Universal Analytics with User ID track that person as a unique user instead of a unique visitor which was previously linked to a browser/device"; and GOOG-BRWN-00711561 at -561: "Enabling the User-ID feature lets Analytics present a cross-platform, cross-device view of your users' behavior…analytics creates a single user journey from all the data that is associated with the same user ID."

[27] See e.g., GOOG-CABR-05454633, GOOG-CABR-04410117, and GOOG-BRWN-00033024.

[28] GOOG-BRWN-00562313 at -316

tracking beacon and saved by a browser, containing identifying data that can subsequently be sent back to the same domain that sent or set the cookie. In effect, cookies allow websites and entities such as Google to store data on the user's computer that helps to identify the user or know something about the user during a sequence of interactions. Web cookies can be valid for a short period of time (a session cookie) or can be persistent for longer periods of time, such as 365 days. Many websites rely on cookies to provide continuity during a web experience. For instance, cookies enable websites to remember that a user has logged in, or to keep track of items in a user's shopping cart. Without cookies, some of the existing web features would not function.

109. Cookies may be sent to a user's browser from Google servers while users are visiting non-Google websites that contain Google's tracking beacons. When the user visits non-Google websites, Google cookies (or the identifiers contained within the cookies) may be sent back to Google along with information from the intercepted communication. Google uses more than [redacted] different types of cookies across its Analytics and advertising products to track and identify users' browsing activities across the Internet,[29] including when users are visiting non-Google websites in private browsing modes without being signed into any Google account. I discuss some of these cookies in Appendix D, along with a description of the impact of browsers enabling cookie blocking.

A.3 Google designed its tracking beacons to collect information by intercepting communications while they were in transit

110. Google designed its tracking beacons to collect information by intercepting communications while they were in transit. When a webpage is in the process of loading, the browser retrieves multiple assets from the web server. During this process, Google tracking

[29] A subset of the cookies Google uses are listed at GOOG-CABR-05435664.

beacons are loaded and used by Google to intercept communications between users and non-Google websites to obtain information that is contemporaneously copied from the initial GET message and sent to Google servers. This occurred continuously throughout the class period.

111. Google instructs non-Google websites to install Google's tracking beacons on their webpages to be executed during the page load process for this purpose, so that while the communication between the user and the non-Google website (the requests to and responses from the website consisting of all the data required for the user's browser to assemble the requested web page) are in transit, Google can intercept the communication contemporaneously. Google requires its tracking beacons to be installed "immediately after the <head> tag on every page of [their] website."[30] Putting tracking beacons in the "head" section that comes first in a webpage (rather than in the "body" section which comes after the "head" section) ensures that the Google tracking beacons start to be executed right away, while the initial user communication is in transit and before a web page loads completely, even if a user starts to navigate away from the page before the page fully loads. Furthermore, putting the Google tracking beacons on every page of a website ensures that users are tracked at every step along the way as they browse a website as soon as they try to load a webpage.

112. In its Answer, "Google admits that . . . [the] Google Analytics scripts . . . provide services as soon as a page is loading." Google's Answer ¶ 74. Google also admits that, "if a user visits a website publisher that uses Google Ad Manager services and includes specific Ad Manager scripts in their website's code, . . . a resource request can be initiated to a Google Ad Manager domain

[30] https://developers.google.com/analytics/devguides/collection/gtagjs, https://developers.google.com/publisher-tag/guides/get-started

https://developers.google.com/publisher-tag/guides/general-best-practices, https://support.google.com/adsense/answer/9274516?hl=en (Last accessed on April 11, 2022).

before the content for the webpage has fully loaded." Google Answer ¶ 80. Similarly, Google admits that its tracking beacons are "retrieved and executed by the browser as part of rendering the publisher's website" (RFA No. 31). "The browser retrieves the embedded code and executes it as part of rendering the website. When executed, this code instructs the user's browser to request content directly from the third-party servers. In making these requests, a copy of the GET request along with certain basic information necessary to provide the service (e.g., IP address, user-agent field, referral URL, applicable cookies, and other information) may be included to facilitate processing the request and providing the requested content. When received, the browser incorporates the requested content into place as necessary to present a seamless webpage." (Google's Response to Interrogatory 40).

A.4 Google's tracking beacons neither facilitate nor are they incidental to communications between users and non-Google websites

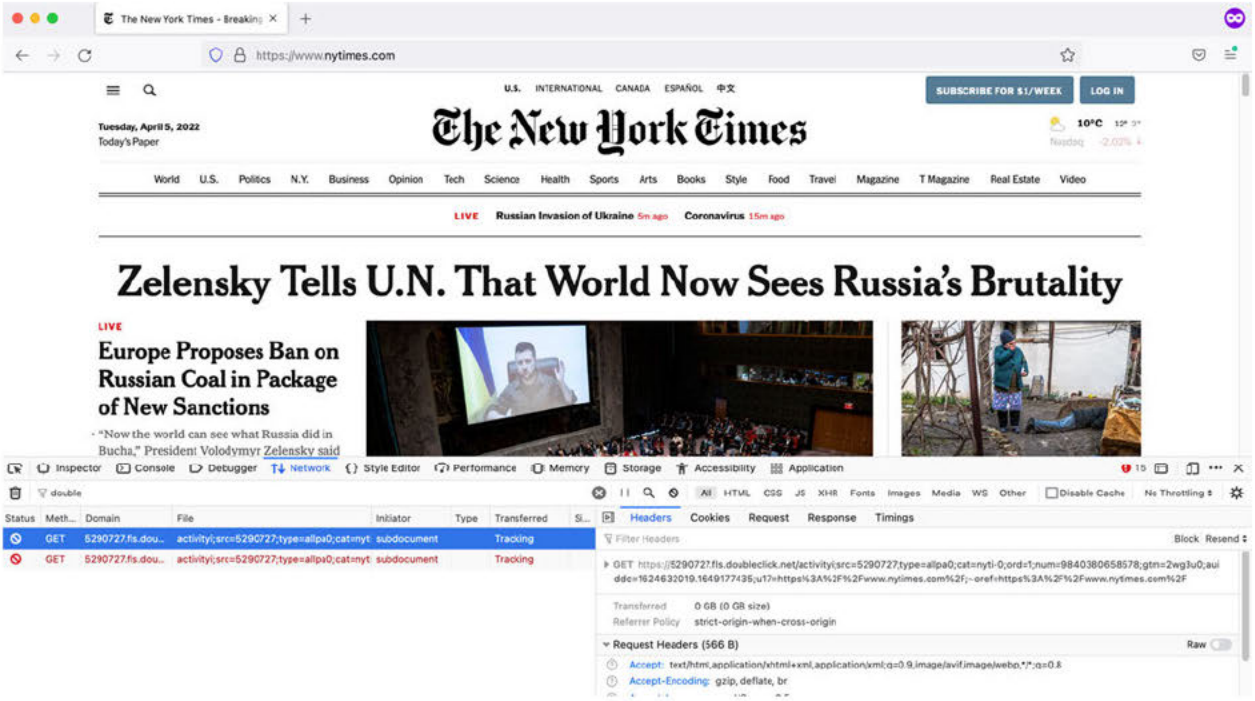
113. Moreover, Google's tracking beacons do not facilitate the transmission of communications between the user's browser and the non-Google websites. The communications flow back and forth between the user's browser and the non-Google website's server irrespective of the Google tracking beacons. Similarly, Google's tracking beacons are not incidental to the user's communications with non-Google websites' servers (i.e., users can communicate with non-Google websites irrespective of the tracking beacons, including if non-Google websites serve advertisements to users, possibly through advertising technology vendors). In fact, without Google's tracking beacons, webpages will load faster and consume less battery energy and bandwidth on the user's device and network. Google tracking beacons harm Plaintiffs and class members by occupying bandwidth on their devices, slowing the performance of the devices, and

cause the devices to consume more power, which is especially significant on battery-powered devices.[31]

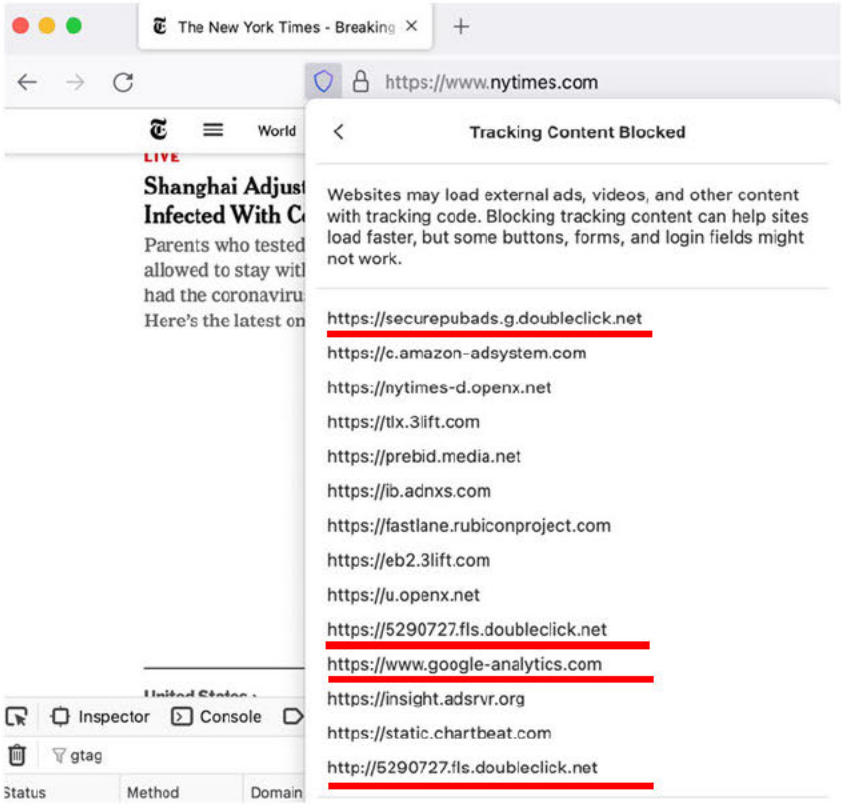
114. My position regarding how these Google tracking beacons do not facilitate and are not incidental to the communications is consistent with the fact that Firefox Private Browsing mode has, since 2015 and by default, blocked Google tracking beacons from causing information to be sent from the user's browser to Google analytics and advertising servers, yet the browser can still communicate with and access webpages.

115. For example, the screenshot below is from Firefox private browsing mode with the bottom half of the screen showing the Firefox Developer Tool. As can be observed, the New York Times page loaded just fine while Firefox blocked the Google tracking beacons attempting to send information to doubleclick.net. In addition, Google Analytics and Google Ad Manager tracking beacons were prevented from operating. Thus, I did not observe any Google ads on the New York Times page when browsing in Firefox private browsing mode.

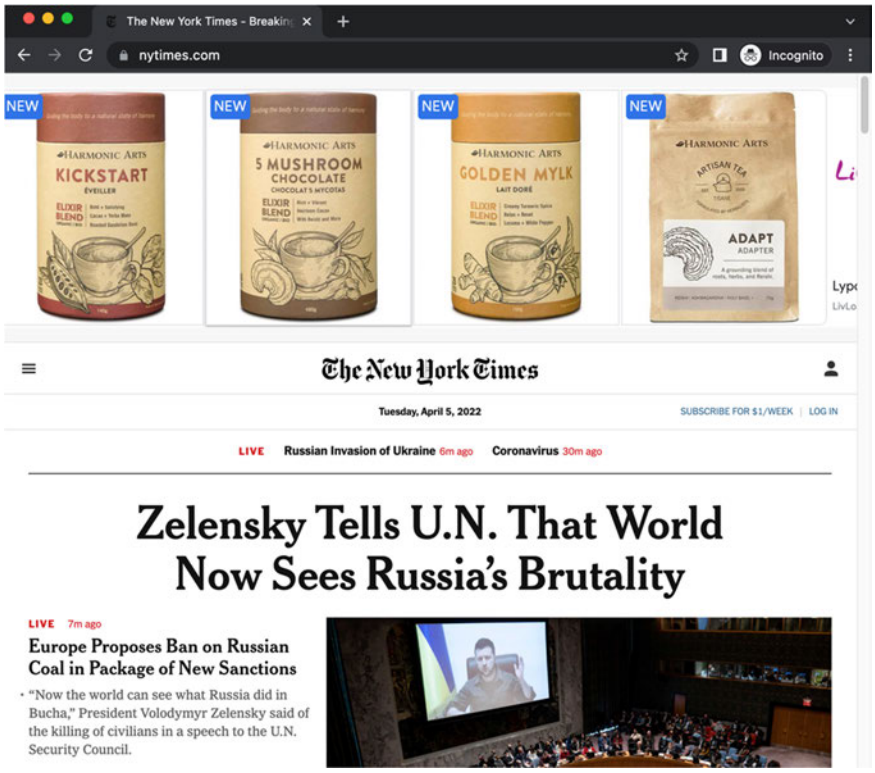
[31] For example, https://blog.mozilla.org/en/products/firefox/firefox-private-browsing-vs-chrome-incognito/ (Last accessed on April 11, 2022) shows Firefox's private browsing with Tracking Protection blocking tracking beacons loads webpages 2.4 times faster than Chrome Incognito mode.



Firefox shows that Google tracking beacons are prevented from sending data to several Google domains as I have underlined in red below.



116. In contrast, Chrome Incognito does not block Google tracking beacons. Thus, the very first thing a user sees when visiting New York Times in Chrome Incognito mode is Google ads right at the top of the page as shown in the screenshot below.



A.5 This functionality was not an accident; Google designed its tracking and advertising code this way

117. Google designed its tracking beacons to intercept communications, even while users were in private browsing modes. This was not an accident, as this was how Google designed this to work. *See* Berntson 30(b)(6) June Tr. 282:11-15 ("Incognito mode was designed so that no system like ads or Google Analytics can know whether or not you're in incognito mode . . ."). Similarly, Brian Rakowski testified that "just like [Incognito mode] doesn't control server logging for any site, it doesn't control server logging for Google" (Rakowski Tr. 141:14-18).

118. Google also designed this process to work such that Google does not offer websites one set of tracking beacons with restricted functions for private browsing mode and another set of tracking beacons for non-private browsing mode. Instead, throughout the class period, Google offered websites just a single set of tracking beacons that perform the same functions whether the user is in private browsing or non-private browsing.[32]

119. Google has admitted in this case that "Google code provided to third party websites for the purpose of providing these services is not designed to differentiate between private browsing/Incognito modes and other browsing modes."[33]

A.6 Google could have designed its products differently

120. Google as a technical matter could have at any point before or during the class period redesigned its various products – including Chrome Incognito – to either stop or limit the information collected by Google by way of Google's tracking beacons embedded within non-Google websites while users were using private browsing modes. The following discusses changes to Chrome Incognito mode first and then changes that would impact Google's collection of private browsing information across the different private browsing modes.

121. Google could have at any point redesigned Chrome Incognito to prevent the Chrome browser from sending information to Google when a user visits a website that uses Google services, such as Google Analytics or Google Ad Manager. *See* Google response to Interrogatory

[32] *See* e.g., Ganem Depo 40:15-24 (testifying that, as Google Analytics product manager, he was not aware of any information Google provided to its customers that would have allowed them to configure tags on their sites to only sent hits to Google when a visitor is not in private browsing mode).

[33] RFA No. 2; *see also* Resp. to Interrogatory 1 ("The fact alone that a user is in private browsing mode does not change the types of data collected when Chrome users visit a non-Google website that uses Google Analytics or Ad Manager while not logged in to their Google Accounts . . ."); Google's Answer, Dkt. 387 ¶ 5 ("Google admits that scripts for Google services that third party publishers choose to include in their websites' code are not designed to differentiate between browser modes.").

29 ("Google has not contended in this case that it is impossible to redesign Chrome Incognito to prevent the Chrome browser from sending information to Google when a user visits a website that uses Google services, such as Google Analytics or Google Ad Manager.").

122. Indeed, since 2015, the private browsing mode for Mozilla's Firefox browser attempts to block ads and analytics tracking beacons by default, including Google tracking beacons. This "Tracking Protection" feature blocks tracking beacons such that Google cannot serve advertisements to Firefox private browsing users who visit non-Google websites with embedded Google tracking beacons.[34] Firefox private browsing also attempts to block Google Analytics and conversion tracking beacons. *See, e.g.*, GOOG-BRWN-00027314 at -314 (internal Google document describing how Google conversion tracking beacons "are all blocked by [Firefox] Private Browsing mode"). Firefox later rolled out Enhanced Tracking Protection (ETP), which includes blocking tracking content in private browsing mode along with third-party cookie blocking and other features.[35]

123. Firefox is an open-source project. Therefore, Google could simply download the Firefox code base and adopt whatever method Firefox has successfully deployed to block tracking beacons in private browsing mode. Instead, both before and throughout the class period, Google chose to allow Chrome Incognito users' private browsing data to be intercepted by Google tracking beacons on non-Google websites. Google at various times considered changes to Chrome Incognito (e.g., GOOG-CABR-04511160 at -160 (discussing proposal for Incognito that would "block[] all 3P trackers AND more")), but Google did not implement those changes.

---

[34] https://blog.mozilla.org/en/products/firefox/firefox-gives-you-more-control-over-your-data-in-private-browsing/ (Last accessed on April 11, 2022).

[35] https://support.mozilla.org/en-US/kb/enhanced-tracking-protection-firefox-desktop (Last accessed on April 11, 2022).

124. Based on my analysis, there were many viable changes to Chrome Incognito mode that Google could have made before the class period – including many changes that Google employees actually proposed but that were never implemented. Those technical changes to Chrome Incognito mode would have limited Google's collection and storage of Incognito browsing information.

125. As one example, while Google finally began blocking third-party cookies by default within Incognito mode in May 2020 (project [redacted], see Appendix D), Google could have as a technical matter implemented this change far earlier. Former Google engineer Justin Schuh admitted that Google could have as a technical matter blocked third-party cookies by default within Incognito mode when Chrome launched in 2008 (Schuh January 6, 2022 Tr. 148:10-20 ("So when you ask me if Google could technically block third-party cookies on – at any point in the history – like, could Chrome had launched without third-party cookies, the answer is technically, yes."). Google employees considered implementing third-party cookie blocking by default far earlier than 2020, but Google ultimately waited until 2020, long after other browsers had already done so.[36]

126. Numerous Google employees within the Chrome group also considered a proposal for Chrome Incognito whereby "Server-side logs will be initially anonymized and then removed at the end of the Incognito session when the user enables this feature." GOOG-BRWN-00164056 at -057. This proposal "was the consensus opinion of our Chrome privacy team in general" and yet Google never implemented it (Kleber March 18, 2022 Tr. 19:6-7-21:6; *see also* Mardini November

[36] *See also* GOOG-BRWN-00410076 at -076 (internal Google email from 2008 noting: The most compatible behavior here is to match Safari exactly. Safari has third party cookie blocking on by default (albeit with a different policy than ours). If I were to optimize for compatibility, I would recommend enabling third-party cookie blocking by default and matching Safari's behavior exactly. I realize that there are other reasons why we're not doing this (pressure from the ads team, for example), but it seems eminently reasonable to allow users to opt in to matching Safari's cookie behavior."); GOOG-BRWN-00397243 at -244 (August 2016 email from Just Schuh stating that "I realize that this is subject is fraught with peril, but maybe it's time to take another crack at a phasing third-party cookie blocking into Chrome by default?").

24, 2021 Tr. 383:21-384:1 (admitting that Chrome "should work on a plan to send [a] signal to Google to delete logs" generated while users are browsing in Incognito mode)).

127. In another example, Google could have the browser set a header in Incognito mode that Google could check to see whether its tracking beacons should run, and if any data transmitted by tracking beacons should be stored. *See* Palmer Tr. 105:8-16 (confirming that the "client-side software can send or not send certain pieces of information, including a do not track header" so it would be "technically possible" for Chrome Incognito to add a Do Not Track Header).

128. Google employees also considered but never implemented an [redacted] which would "introduce anti-tracking features—including invasive anti-fingerprinting measures" (GOOG-CABR-05269357.R at -79); *see also id.* at -91 (defining "tracking" as "third-parties following a user's browsing behavior").[37] Google employee Michael Kleber explained that [redacted] (Kleber March 18, 2022 Tr. 23:16-24:11). "[E]verybody involved in the discussion was in favor of figuring out ways in which we could provide this sort of protection" and yet Google never implemented that [redacted] (*id.* 27:21-25).

129. Google also considered but never implemented a proposal to [redacted] (GOOG-BRWN-00650016 at -16). The goal of this proposal was [redacted]

[37] Google relatedly considered but never implemented an [redacted] that would [redacted] (GOOG-BRWN-00155943 at -943).

[redacted]

(Kleber March 18, 2022 Tr. 43:20-44:4, 47:2-11). Internal Google documents also describe how adopting this proposal would mean that [redacted] (GOOG-BRWN-00650016 at -016). Mr. Kleber also explained that "I think everybody [within Chrome] is in favor of IP privacy" (Kleber March 18, 2022 Tr. 51:8-10). But unlike Safari, which has already implemented an "IP blinding system," Chrome has not yet introduced such protections (*id.* 54:4-15).

130. Google has also considered but never implemented a different kind of Incognito that would not allow what Google refers to as "session-based" tracking. Instead of multiple Incognito windows . . . "hav[ing] [a] shared state, in that what happens in one of them might influence what happens in another one later," "two different Incognito windows inside the same browser [would] behave[] independently of each other and more like Incognito windows in two different unrelated Chrome browsers" (Kleber March 18, 2022 Tr. 56:17-58:6; GOOG-BRWN-00555073 at -084). With Safari's private browsing mode, "browsing initiated in one tab is isolated from browsing initiated in other tabs, so websites can't track browsing across multiple sessions."[38] It is my opinion that Google could revise Chrome Incognito to function in the same way, and Google employees have considered such a change. But Google has not implemented this change.

131. Furthermore, Google could have before or during the class period built a toggle on the Incognito screen (sometimes referred to by Google employees as the New Tab Page or NTP) giving users "the option to opt in to signal Google and only Google of their Incognito browsing sessions, so that on the Google web services, the web server could be aware of that . . . session being an Incognito one and, therefore, could treat that session's data differently, either through not

[38] https://www.apple.com/safari/docs/Safari_White_Paper_Nov_2019.pdf

capturing it in the first place or limiting it to lifespan" (McClelland Tr. 87:14-89:6).[39] Google had even built a browser meter for selected users called Panelists. This browser meter was able to recognize Incognito mode and prevent data collection in Incognito mode.[40] But Google never implemented this approach to Incognito more generally.

132. Google could have also made changes to its tracking beacons with respect to how those tracking beacons would operate on other private browsing modes. For example, Google could have respected a Do Not Track protocol ("DNT") analogous to the Robot Exclusion Protocol. Chrome is and has been since at least 2016 the leading browser by market share,[41] which places Google in a position of industry leadership. But Google instead chose not to honor such a signal because it would hurt Google's revenues (McClelland Tr. 126:14-127:21, 131:12-132:13). Additionally, a Google tracking beacon using JavaScript could place a message on the user's screen within the browser window to request permission for collecting and sending data to Google.

133. Finally, Google could have described its interception and collection of user data on non-Google websites while the user is in private browsing mode. Google employees recognized this option, including with a proposal to design an icon that makes clear "You are not protected from

[39] Google's internal documents suggest that the Chrome team went so far as to consider implementing such a change. *See e.g.,* Halavati Depo. 98: 9-16 ("I worked on the possibility of telling Google that user is in more private browsing mode. . . We had a very long discussion about it, more than a year"); *see also id.* at 102 11-14 ("So assuming that Google services log things when users is in Incognito because they don't know user is in Incognito, if we had told Google that user is in Incognito, they could be more protective about it.").

[40] "Browser meter is an extension installed on a browser. It collects browsing activities and cookies from the browser after panelist signing into the meter. It does not collect data in incognito mode or in "guest mode" (used by a guest who is not a panelist)." (GOOG-CABR-04073287 at -291)

[41] https://gs.statcounter.com/browser-market-share/all/united-states-of-america#monthly-201601-202202 (Last accessed on April 11, 2022).

Google." (GOOG-BRWN-00048773); *see also infra,* Section K, for more examples of how Google could have described its interception and collection of private browsing information.

B. User Notification & Choice: Throughout class period, Google collected this private browsing information without notifying users or offering a choice at the time of or in connection with any of the interceptions

134. When visiting a non-Google website containing Google tracking beacons in a private browsing mode, users are not notified of the Google tracking beacons or Google's collection of private browsing information or given a choice regarding that collection. I visited each of the top 25 websites[42] identified by Google for Google Analytics and for Google Ad Manager as well as other websites outlined in Appendix C in private browsing modes, including Chrome Incognito, and none of those websites had any such pop-up notification – or any process by which users would be informed that Google would collect and exploit their private browsing information or any option to prevent such collection. Google's internal documentation, when discussing that Google wants to say "You have full control over the data Google collects about you. You may dictate how Google can use that information about you, and at any time can stop Google from collecting that data about you", acknowledges that "From a Product perspective…We would not be able to say this because we will collect AdID, IP address, browser headers, searches etc regardless of setting (both signed in and signed out). We *Collect* this data regardless of the user settings" (GOOG-BRWN-00466617 at -637).

135. Based on my analysis, Google has systematically failed to provide users in the United States with any such notification of Google's interception, collection and use of private browsing information when they visit non-Google websites, and that has been consistent throughout the class

[42] Google, in its SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 5, has identified howto.gov as one of the 25 top-level domains registered in U.S. time zones with the highest traffic for Google Analytics. However, this domain does not appear to exist on the Internet.

period, across class members, and across all websites containing these Google tracking beacons. These Google tracking beacons do not announce themselves to the user or allow the user to choose whether to send the user's private browsing data to Google, and then to be exploited by Google. Google could have included either a pop-up notification or a choice to users at the time collection, but it did not. The developer tools (such as the Chrome Dev Tools) that I used to conduct my analysis are not designed for use by typical Internet users, and they in any case do not provide information regarding exactly what private browsing information is collected by Google, what private browsing information is stored by Google, and/or how that private browsing information is exploited by Google. Google does not enable users to turn off its data vacuum in private browsing modes such as Chrome Incognito.

C. Website Notification & Choice: Throughout class period, Google collected this private browsing information without notifying websites at the time of or in connection with any of the interceptions or providing websites with a choice

136. Google does not provide non-Google websites with Google tracking beacons any option to turn off interception and collection in private browsing mode on any browser, including the Chrome browser, including prior to the interception. I am a Google Analytics and Google Ads customer, and Google has not provided me with an option to ensure that its tracking and advertising code would avoid collecting data from private browsing sessions.

137. While Google sends itself Incognito usage signals when users browse in Incognito mode (see Section H and Appendix G), Google prevents any other website from being able to distinguish Incognito and regular Chrome browsing communications. Google provides tracking beacons to non-Google websites on a take it or leave it, all or nothing basis with regard to Google's collection

of private browsing information, while telling websites that Google respects privacy.[43] Google does not enable non-Google website operators to selectively turn off Google's data vacuum for users in private browsing modes such as Chrome Incognito, and that has been Google's practice continuously throughout the class period and across websites.

D. Google Storage: Throughout the class period, Google stored private browsing information in many Google data sources, sometimes permanently

138. Based on my analysis in this case, including in connection with the data-focused discovery and productions by Google both through the Special Master process and otherwise, I conclude that Google stores private browsing information collected from users' private browsing activities, and that Google has done this systematically throughout the class period and across the classes when users are visiting non-Google websites. Google stores private browsing information in many Google data sources (including Google logs), sometimes permanently. The following provides additional information regarding this Google storage.

139. Google's internal document acknowledges that "Display ads and its protos store a wealth of information about users (visiting advertiser or publisher web page urls). This can be explicit i.e. can be in the form of cookies (biscotti/zwieback/gaia etc), hardware (android log id) or software id (user agent), location, ip addresses etc." (GOOG-CABR-04141174). Browsing in private browsing mode does not prevent Google from logging intercepted and collected private browsing information. As Google's own documents state, "Incognito also does not prevent server-side

[43] *See* https://policies.google.com/technologies/partner-sites?hl=en-US ("Google uses the information shared by sites and apps to deliver our services, maintain and improve them, develop new services, measure the effectiveness of advertising, protect against fraud and abuse, and personalize content and ads you see on Google and on our partners' sites and apps. ***See our Privacy Policy*** to learn more about how we process data for each of these purposes and our Advertising page for more about Google ads, how your information is used in the context of advertising, and how long Google stores this information." (emphasis added)).

logging" (GOOG-CABR-05149122 at 124); *see also* GOOG-BRWN-00153850.C. at -856 ("'Just for you sounds like Google doesn't log your activity in Incognito – but we do."); Google's Response to Request for Admission 43 ("Google stores the data at issue in this case in Analytics and Ad Manager protocol buffer logs . . .").

D.1 With the Special Master process, Google identified some but not all sources that include private browsing information

140. Through the confidential discovery in this case, I was able to investigate and understand the processes by which Google collects, stores, and uses private browsing information. This discovery process was in part supervised by a Special Master, who assisted with those efforts in terms of identifying and producing information regarding Google's logging of information, including private browsing information. I was directly involved in these efforts, including by evaluating the information and data produced by Google. Those efforts were a core part of my analysis in terms of developing opinions regarding Google's collection, storage, and use of the private browsing information at issue in this lawsuit, for the class and across the class period.

141. Throughout the class period, Google collected the intercepted private communication information and stored them in various Google data sources, including many Google logs. On September 17, 2021, pursuant to a Court order, Google identified [redacted] logs as "potentially relevant sources of ESI [Electronically Stored Information]," including a short description for these sources. On September 24, 2021, again pursuant to Court order, Google provided a list of schema / fields for some of those [redacted] data sources (limiting the fields to the top 100 fields for certain logs). On November 22, 2021, Google identified [redacted] additional data sources and provided lists of schema / fields for some additional data sources within the initial [redacted] provided on September 24, 2021.

142. Subsequently, through discovery, Plaintiffs identified additional data sources not included in the [redacted] data sources identified by Google. This included a Google data source named [redacted], which Google engineers used to perform log-level Incognito detection of Chrome traffic, right around the same time that this lawsuit was filed (GOOG-CABR-05280756); *see also* Google's Supplemental Response to Interrogatory No. 34 (identifying this log as one of the logs studied for Google's Incognito-detection analysis). Google provided the top 100 fields for [redacted] on December 16, 2021. Belatedly, on March 17, 2022, Google identified for the first time [redacted] other logs that Google engineers considered in their log-level based Incognito detection study (Google's Supplemental Response to Interrogatory No. 34). These Google logs had not previously been identified to the Special Master.

143. Two weeks before discovery closed, in response to requests from Plaintiffs for Google to produce Google engineer Bert Leung's documents, Google produced documents indicating that Google is actively logging an Incognito detection field called "maybe_chrome_incognito" in certain Google ads logs.[44]

144. On or about March 11, 2021, Google employees drafted a logging proto to implement this "maybe_chrome_incognito" field (Stipulation Regarding Google's 'Maybe_Chrome_Incognito' Field). Google began logging this field in nineteen specific Google logs by July 4, 2021 (*Id.*)

145. This "maybe_chrome_incognito" field contains two values: TRUE and FALSE. It is FALSE when an event was determined by Google to be *not "*maybe_chrome_incognito" and TRUE when an event was determined by Google to be "maybe_chrome_incognito" (Liu Tr. 20:6-8). Through logging this field, Google has since July 2021 been tracking user browsing activities

[44] GOOG-BRWN-00845312 at -318: "Bert Leung, 2021-09-13 14:58:47: are we already logging maybe-chrome-incognito in search [redacted] already? Mandy Liu, 2021-09-13 14:58:54: yes"

in Chrome Incognito mode, designating those activities as "maybe_chrome_incognito" for Google's own internal purposes - but without any efforts by Google to use that field to stop Google's collection, logging, and use of any private browsing information.

146. On March 1, 2022, in response to requests from Plaintiffs, Google provided the names of the [redacted] Google log sources containing the "maybe_chrome_incognito" field.[45] Only [redacted] of these log sources ([redacted] and [redacted]) had been previously identified by Google, with the other [redacted] Google log sources having been omitted from Google's prior disclosures. And for the [redacted] logs that were identified, Google omitted the "maybe_chrome_incognito" field from the schema that was produced to the Special Master and Plaintiffs.[46]

147. With my involvement in this Special Master process, Plaintiffs would have identified these [redacted] log sources as important logs to search for Plaintiffs' data months earlier had Google produced full schema for these logs, but Google instead produced just the (purportedly) largest 100 fields – leaving out the "maybe_chrome_incognito" field. Google could have produced more fields for these [redacted] logs, as Google did for other logs (e.g. [redacted] fields for [redacted], [redacted] fields for [redacted], and [redacted] fields for DBL [redacted][47]). Even if somehow the top 100 fields were easier to provide, Google could have at a minimum added the "maybe_chrome_incognito" field in the schema, given its significance, but Google did not. Thus, for months we were led down different paths, searching other logs while Google kept hidden this "maybe_chrome_incognito" field.

[45] 2022-03-01 Brown v. Google - Google Letter re [redacted] Monitoring Dashboard.pdf

[46] Brown - 3.8-3.9 Field Names

[47] 2021-11-18 Brown Omnibus Sheet.xlsx and 2021-12-06 Brown - Special Master Submission.xlsx

148. On March 4, 2022, when asked to produce the schema for the [redacted] log sources containing the "maybe_chrome_incognito" field, Google continued to produce just the purportedly largest 100 fields for these log sources, leaving out the "maybe_chrome_incognito" field in all but one of the [redacted] logs.

149. On March 11, 2022, for each of the [redacted] logs containing the "maybe_chrome_incognito" field as well as [redacted] ads logs and one Analytics log, Google produced additional schema / fields pursuant to an order by the Special Master. This time, the schema produced showed the presence of the "maybe_chrome_incognito" field in [redacted] of the [redacted] logs.[48] In addition, this production revealed the presence of an "is_chrome_non_incognito_mode" field in [redacted] of the logs that Google engineers used for their log-based Incognito detection study. *See* Google's Supplemental Response to Interrogatory No. 34. This additional Incognito detection field was omitted from the prior schema production for these [redacted] logs.

150. The logs containing "maybe_chrome_incognito" were not the only logs that Google failed to identify. Google logs such as the Chrome Sync logs and [redacted] logs containing an "is_chrome_non_incognito_mode" field were also omitted by Google. I first learned about this field in Chrome Sync logs in connection with a Special Master live demo session on March 4, 2022, attended by two of the consultants I worked with in connection with this report (Dr. Dai and Mr. Thompson). I understand that Google still has not identified all logs containing this field. Based on my review of the documents produced by Google, it remains unclear how Google assigns a value to this field and how it is used by Google. The only reference to the existence of this field I found in Google's productions is in [redacted] logs (GOOG-BRWN-00176433), which is a Google

[48] Because the schema / fields were generated through data searching using just three identifiers, the schema thus generated is not a full set of schema contained in each log.

[redacted] (GOOG-BRWN-00536949). Google also produced User Metrics Analysis ("UMA") data containing this field in January 2022, but Google has not yet provided any explanation as to how this field relates to UMA data.[49]

151. On March 11, 2022, Google engineer Dr. Sadowski for the first time disclosed the existence of [redacted] logs – [redacted] containing an "is_chrome_non_incognito" field and [redacted] containing an "is_chrome_incognito" field.[50] Dr. Sadowski testified that these Incognito fields have been in use by Google in [redacted] logs since 2017 and are still in active use (Sadowski Tr. 70:11-19). On March 31, 2022, Google produced the top 100 fields of the five identified [redacted] logs, which, once again, left out the "is_chrome_incognito" field. The "is_chrome_non_incognito_mode" field appeared in three logs along with user location information.[51]

152. On April 1, 2022, Google provided the proto comment for these fields.[52] For the "is_chrome_incognito" field, the proto comment states [redacted] and for the "is_chrome_non_incognito_field", the proto comments states [redacted] Google has not produced [redacted] in its document production.

153. Google's internal documentation reveals many other logs that may contain private browsing information. Other logs Google withheld from discovery are included in Appendix E.

[49] Google's counsel wrote a letter in response to questions Plaintiffs raised about this "is_chrome_non_incognito_mode" field. But Google failed to provide any explanation about why or how this field appeared in the raw UMA log. *See* 2022-03-04 SM Letter Follow Up - Google edits.pdf

[50] Final Sadowski Reference Sheet.pdf

[51] GOOG-BRWN-00848368

[52] 2022-04-01 Brown v. Google - Google Letter.pdf

D.2 Google keeps some private browsing information permanently, with no option for users to review or request deletion of that information

154. Based on my investigation in connection with the Special Master process, I understand that Google keeps some of the information it collects from users' private browsing activities permanently (including ad queries, ad views and conversion data)[53], with no option for users to review or request deletion of that private browsing information.

155. Google offers users the ability to delete certain data from their browsing history, including URLs and page names.[54] However, this control is only offered to logged-in Google users, and does not affect data stored while a user is in a private browsing mode while not logged-into Google. In any event, when a user deletes activity from their Google account, they will see the activity as deleted from their Google account; however, the data is not immediately purged from Google's logs, which may take up to [redacted] for deletion.[55]

156. As for the private browsing information at issue in this case – where class members are not signed into any Google account – Google stores the data in a way that people have no ability to review or delete that information – with various identifying information that could be linked to class members and their devices. Google provides no process to review and delete that private browsing information from Google's servers – or preventing Google's use. Even Google's Senior Software Engineer for Chrome testified that he did not think Google provides users an option to

[53] Data Source Information Provided Reconciliation - Brown.xlsx; *see also* Google's Response to RFA No. 46 ("Google further admits that some of this data [at issue in this case] is retained permanently.").

[54] https://support.google.com/accounts/answer/7028918 (Last accessed on April 11, 2022).

[55] GOOG-BRWN-00432838 at -840: "user activity history deleted from [redacted] UI …The corresponding information is tombstoned in PLogs in the [redacted] windows". Google engineers at "Commented[7]" asks why deleting data in logs can't be instant as well. "If we tell users that we will give them more control on their ads experience by deleting their history. They may have an expectation of immediately seeing their targeting change after a deletion."

prevent Google logging while they are in Incognito or private browsing mode (Halavati Depo. 101:15-25); *see also* McClelland Tr. 320:8-14 ("The My Accounts section is only there to control authenticated data. Q. So a user does not have a way to delete unauthenticated data from Google logs, correct? . . . A. That is my understanding.").

157. Internal Google documents also confirm this. For example, as one Google engineer explained, "As a signed-in user, I have tools to see what data Google has collected on me, selectively delete data items, and even 'take it out'. As a signed-out user, we provide nothing like this" (GOOG-BRWN-00433503 at -504); *see also* Berntson March 18, 2022 Tr. 171:17-19 ("So to answer your question, is there a way for a user to go back and view the data associated with an Incognito session, no, because all of the IDs have been deleted when they close the Incognito session."); GOOG-BRWN-00153850.C. at -56 ("Incognito mode (when you haven't signed in), your history is stored to a non-account identifier and you don't have the ability to see or delete it"); Fair Tr. 209:8-20, 210:8-18 (admitting that logged-in users have "more robust" control over their data than logged-out users).

158. Because there is no way for a user to subsequently exercise control over the private browsing information collected by Google (either to view it or request deletion), in effect, Google provides signed-out users (including all class members in this case) with *less* control over their privacy in the private browsing mode than when they are logged into Google in normal browsing mode.

159. Google could have offered users the ability to have Google delete their Incognito browsing information while signed-out of their Google account. Google employees recognized that providing that option would be a way for Google to make good on a "key promise" to Incognito users that Google does not "collect data from the Incognito session". GOOG-BRWN-00147864 at

–865 and at –866 (recognizing offering a signal at the end of the Incognito session could give Google a way to "de-identify/delete/anonymize all Incognito session data"). In GOOG-CABR-05881692 at –692 and at -693, Google describes an architecture where "signed-out user can request 'view' and 'deletion'. Sign-in is not required." This includes a "new browser UI [user interface]" where "the user can instruct all ad tech companies [such as Google] to delete all user data, or user data in selected categories."

160. Google distinguishes between signed-in and signed-out modes by using different cookies (GOOG-CABR-03716577 at -586), and Google stores the intercepted data in different logs depending on the signed-in vs. signed-out mode. Google uses GAIA cookies for signed-in mode and what Google calls "pseudonymous" cookies such as Zwieback and Biscotti cookies for signed-out mode (which, in fact, can be linked to users through fingerprinting information Google collects in its logs, such as IP address, user agents, websites visited, and more). In GAIA Mode (also called signed-in mode), Google logs data to what it refers to as personal logs (P logs); in signed-out mode, Google logs data to Biscotti logs (B logs) for display ads data and Zwieback logs for search ads data (though that is not limited to just google.com searching).

161. The signed-out mode represents a great majority of the use cases for private browsing usage. When a user starts a new private browsing session, they are automatically signed-out of any signed-in Google account and most users do not then sign into Google in the private browsing session.[56] One Google engineer explains, for example, "Users in incognito mode usually don't

[56] GOOG-BRWN-00697719 at -719: "If you're browsing in Chrome Incognito mode, you are, by default, not signed into any accounts or sites", GOOG-BRWN-00027102 at -103: "GAIA identifier is not used in incognito mode unless the user explicitly logs in", "Safari's Private Browsing mode creates a completely new session for each private tab — or "new private window" — you open. These temporary private sessions are completely isolated from one another as well as from your primary browsing window. That means: You won't be signed in to any of your accounts…" (https://www.avast.com/c-how-to-go-

sign-in with gaia, thus their clicks/conversions are all sign-out" (GOOG-CABR-04959606 at -609). Other documents indicate that Google sign-in represents a small fraction of traffic in private browsing mode (e.g., GOOG-BRWN-00148029 at -050 indicating that [redacted] of Incognito search queries are logged-out). Unlike signed-in mode, a user has no ability to review or actively delete any of their signed-out mode data.[57]

D.3 Google merged private and non-private browsing information in its logs and other data sources

162. Throughout the class period, Chrome Incognito mode did not sandbox individual windows and tabs. Instead, Google designed Chrome Incognito to function in a way where cookies were shared across all Incognito windows and tabs open at the same time.[58] This "means that if you open an Incognito tab inside of Chrome and then you open a second Incognito window inside of Chrome while the first one is still open, then the relationship between those two Chrome windows is the normal relationship between two different Chrome windows in the same browser. That is,

---

incognito-safari Last accessed on April 11, 2022). I have observed the same signed-out behavior on Microsoft Edge's InPrivate mode as well as on Incognito and Safari Private Browsing.

[57] Even in the scenario where a publisher website demands Google to delete data associated with a particular PPID, Google appears to merely delete the link association between the PPID and mapped Biscotti, leaving the actual private browsing data stored in logs intact (GOOG-CABR-04696292 at -301): "PPID based user data in [redacted] and [redacted] records can be ignored once the PPID to Biscotti mapping is broken". (This is a working draft document, but the quotation is not contradicted by anything else I have seen, and this behavior was "Confirmed with legal". See *Id.* at FN 9.)

[58] GOOG-BRWN-00663644 at -701: "incognito sessions are shared among tabs and windows…Many users are unaware of session based tracking" and at -703: "We can isolate incognito sessions per tab, instead of sharing the cookie jar across Incognito tabs and windows

Benefits:

- Isolated cookie jars, prevents the need to educate users about the importance of their session length

- No ad personalization across tabs/windows

- Follows the trend of optimizing for privacy over convenient browsing set by Safari and Firefox"

it's possible for what happens in one of those windows to potentially influence what happens in -- later on in the other window" (Kleber March 18, 2022 Tr. 56:17-25). Accordingly, if there are two open Incognito windows and in each window there are three tabs, all six tabs are sharing the same state (*Id.* 57:12-16).

163. By contrast, Safari's Private Browsing mode was designed to *not* share cookies across all open windows and tabs.[59]

164. Google's design choice has had implications in terms of what data Google logs versus user privacy notifications and choices. Google employees wrote, for example "Most users are not aware of session-based tracking" (GOOG-BRWN-00042388). Another internal Google document notes "'Sign out' is an ambiguous, loaded term that generally means only a portion of what users probably want. Cookies span signed in and signed out sessions, so Google and 3rd-party products can connect the dots even if they can't write data to a person's account" (GOOG-BRWN-00060463). The result is Google logging browsing data on non-Google websites from private-browsing and non-private browsing sessions within the same GAIA logs.[60]

165. In addition, Google stores a user's logged-in identifier on non-Google websites (e.g., Publisher Provided ID ("PPID"), Analytics User ID and Google Ads User ID) in its logs. PPID is mapped to a PPID-mapped-Biscotti ID, which is used to key user data.[61] Whenever a user logs-in

[59] GOOG-CABR-04508088 at -088: "even IF a user understands what a session (within a window) is, many users are still not aware that they need to close the full window in order to reset the session. This approach is a solution because every tab is isolated, therefore more private. (Safari functions like this)"

[60] GOOG-BRWN-00072051 at -052 and -053: "if you sign in to your Google Account while in the Incognito session: Data associated with the temporary ID gets saved to your Google Account; Your additional activity in this Incognito session also gets saved to your Google Account. In other words, if you sign in to your Google Account from within an Incognito session, from Google's perspective, it's as if you aren't in Incognito mode anymore."

[61] GOOG-CABR-04705124 at -124: "1. PPID is a publisher owned and provided identifier. 2. User profiles generated using PPID are scoped to a single publisher network only… We can use PPID for

on non-Google websites, whether in private browsing or non-private browsing mode, the same identifier is associated with the data Google collects from a user's browsing activities on that website. Google further logs all such data (private and non-private) within the same logs and uses these data for serving personalized ads.[62]

E. Google Use: Throughout the class period, Google exploited private browsing information to generate revenue for Google by creating profiles, serving ads, tracking conversions, and in other ways that benefited Google

166. Based on my review of the information and data obtained through the Special Master process, internal Google documentation, deposition testimony, and written discovery, it is my opinion that Google, throughout the class period, collected and kept a massive amount of information about its users, with Google reportedly recording [redacted] of [log] entries a day,"[63] including private browsing information Google collected during the class period by intercepting users' private browsing communications with non-Google websites while those users were not

---

features like f-capping and personalized advertising. By personalized advertising, we mean targeting based on demographics, IBA [Interest Based Advertising] and similar audiences, and publisher's audience profiles… We would use the traversal history of the user on the publisher's network of sites to create an IBA profile"; GOOG-BRWN-00149849 at -852: "The PPID is mapped to a new biscotti ID the first time we see an ad request where the publisher specified a PPID. The PPID is hashed and stored in a table mapping PPIDs to biscotti IDs"; and Liao December 3, 2021 Tr. - 55: 12-24 ("[redacted], I believe, sends the mapped PPID—sorry the PPID mapped Biscottis ID, along with other pieces of data, onwards to the rest of display serving system, including [redacted] . . . [redacted], along with its own back ends, collectively are responsible for some parts of ad serving."); *id.* at 57: 10-20 ("under some circumstances, [redacted] can log the information, including this mapped PPID. . . There are many different logs that the [redacted] writes, so I'm not aware of all of them, but one of the different types of logs in this case will be ad query logs.")

[62] GOOG-BRWN-00149849 at -849: "Publisher provided identifiers allow the publisher to send DFP an identifier to be used in place of the DoubleClick cookie. The advantage in this case is that the publisher may have a consistent identifier across different devices" and at -852: "Any time a PPID is specified in the request, the ad selection will consider the audience segments associated with the mapped biscotti ID and DoubleClick cookie if the user also has a separate DoubleClick cookie on the browser". GOOG-CABR-04122554 at -575 shows ad personalization with PPID recovers [redacted] of ad revenue loss due to Apple's Intelligent Tracking Prevention (ITP) on Safari browsers.

[63] Dkt. 119 at 3.

signed into Google. Google used and continues to use this information for a variety of purposes, including creating detailed profiles about the users, serving ads to users (including targeted and personalized ads), tracking and modeling conversions from that user activity to enhance its advertising revenues, providing analytics information to customers based on user activities, and for algorithm development and training, reporting, process improvement and more.

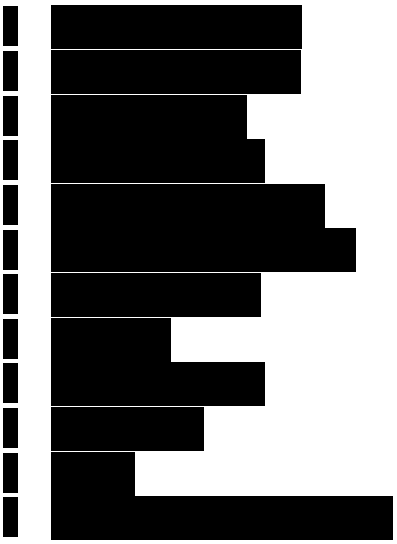
E.1 Google Tracking & Profiling: Throughout the class period, Google tracked private browsing communications to create detailed profiles

167. To deliver targeted advertising to users, throughout the class period, Google built and used detailed profiles tied to each user's identifiers, including when the user is in private browsing mode. Based on my analysis in connection with this litigation, I now understand that Google has done this even when users are privately browsing non-Google websites and when the user is not signed into Google, using various undisclosed Google identifiers to track and target users.

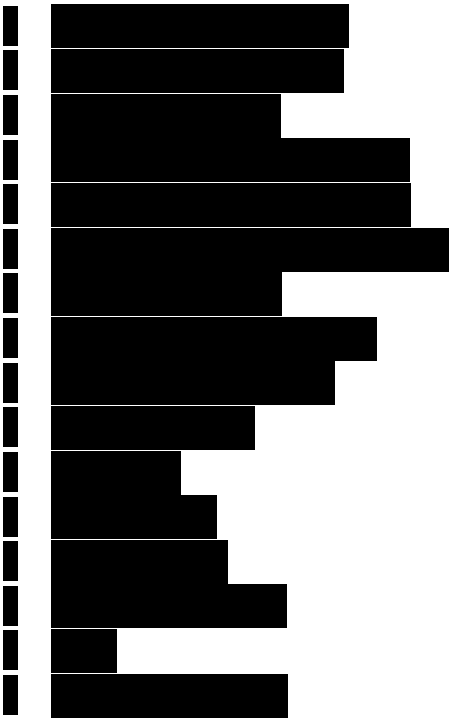
168. Throughout the class period, Google created and then benefited from its use of various user profiles, including while users were in private browsing mode visiting non-Google websites and not signed into Google.[64] Google explains that "data collection is extremely valuable because it allows advertising businesses to create a profile based on your browsing history. The more data these businesses have, the more accurately they can target you with advertisements, which (indirectly) means higher revenue for the company" (GOOG-CABR-04006287 at -287). Based on users' browsing activities, Google creates user profiles based on (GOOG-CABR-05864545):

[64] *See* Halavati Depo. 113:18-25 (confirming Chrome Incognito mode's implementation creates off-the record profiles of users and has since at least August 2016 when testifying engineer started working on Chrome); GOOG-BRWN-00207020 at -23 (describing scenario in which significant other sees "targeted ads for engagement rings or weddings" following Incognito searches for engagement rights, and explaining: "Unfortunately Incognito mode doesn't prevent this awkward moment. Since IP address is a common signal and Incognito sessions build zwieback based profiles, this can, and as has been suggested by some users, and does happen").

- The user's interests or hobbies (what Google refers to as the user's "affinity") – Google's internal document (GOOG-BRWN-00229628) shows [redacted] interest categories. These include the following top-level categories:



- What the user is actively searching/browsing with an intent to purchase (what Google refers to as "in-market") – Google's internal document (GOOG-BRWN-00229632) shows [redacted] in-market categories. These include the following top-level categories:



- User "Demographics" (e.g., the user's age, gender, parental status – inferred based on user's browsing history on non-Google websites) (GOOG-CABR-03742530). Thus, even if users do not provide demographics information, Google can "estimate it from the user behavior" using machine learning algorithms (GOOG-BRWN-00535100 at -134).
- User "Life Events" (e.g., whether the user is graduating from college, moving homes, or getting married).
- Custom Intent Segments - Auto-created custom user intent segment for targeting a specific vertical or landing page.
- User past interactions with a website.
- User remarketing – people who have previously visited an advertiser's website.
- Similar Audience – people who are similar to those on the remarketing list, generated using Google's machine learning algorithm.[65]

[65] GOOG-CABR-04626237 at -238: "Similar Audience can be considered as the expansion of Remarketing. Remarketing is the work of retargeting a group of users, called seedlist, who have visited a site before. Similar Audience is the work to find and target users who are similar to ones in the seedlist. We started the Similar Audience project for GDA in 2009 and later for DV3 in 2014. The core part of the project is using machine learning algorithms to model or learn the similarity between display audiences."

169. There may be several Google user profiles associated with a user, including GAIA-keyed user data and Biscotti-keyed user data. These detailed profiles are core to Google's business, and Google uses these user profiles for ad targeting, frequency capping, and logging.[66]

170. A significant amount of user data, including data collected by Google during the class period from users' private browsing on non-Google websites while not signed into Google, are stored in (among other places) two Google data repositories called [redacted] and [redacted].[67] These are enormous Google data repositories, which Google used during the class period to store user data, including commingled data collected by Google from users' private and non-private browsing activities. Google explains that Google's "DBL" logs in both of these Google storage locations (DBL [redacted] and DBL [redacted]) contain user data "keyed to unauthenticated IDs" that Google uses with "various display ads products for targeting, personalization, user controls and measurement like use cases."[68] The following table includes information provided by Google regarding these two Google data sources:

[66] GOOG-CABR-04695672 at -673: "we look up user profiles using Biscotti id as a key and then use the data for targeting, frequency capping, and logging." Google states that this is a "grossly oversimplified view," but their caveats discuss policy limitations on what data and identifiers can be used when. Policy limitations are not an assurance of privacy because policy can change ("Don't be evil" was phased out), threat actors do not follow policy, and government or legal intervention is not subject to corporate policy.

[67] Google has been moving some [redacted] data over to [redacted] (See e.g., GOOG-CABR-04207454).

[68] "Data Source Information Provided Reconciliation - Brown.xlsx"

171. "DBL" stands for DoubleClick. In these DBL logs, Google stores data in tables, with rows and columns. Each column in the DBL table is keyed by a user's Biscotti ID, which is a unique identifier generated by Google in connection with a user's browsing activities.[69] The data Google associates with a particular Biscotti ID represents data Google collected during the particular user browsing session, where that Biscotti ID is generated and used to track the user's browsing activities. The Biscotti ID generated by Google can be from various sources, all tied to the user browsing activity, including the IDE cookie, PPID-mapped-Biscotti, GFP-mapped-Biscotti or other Biscotti ID sources (GOOG-CABR-04423260 at -269). Google generates and uses these unique Biscotti IDs for the purpose of tracking and profiling users' browsing activities, including without limitation the private browsing activities of Google account holders visiting non-Google websites while not signed into Google, and does this for Google's own benefit.

[69] GOOG-CABR-00892264 at -265: "The Biscotti Id is a number that uniquely identifies a browser or other device. Browsing history and other activities associated with a given Biscotti Id are aggregated over time, allowing models and 'user profiles' to be built… the Biscotti Id is extracted from the Biscotti cookie and used as the key to lookup the user profile in [redacted] and the user profile is then used to help select the best ad." And GOOG-CABR-04423260 at -266: "We use [redacted] to store user data … Different [redacted] 'tables' for different key-spaces (e.g. logged in data, vs doubleclick data). We use the 'dbl' table for doubleclick data. Row key is 'biscotti id', columns represent different data about a single cookie"

172. Google's DBL [redacted] data repository also stores data in a table format, with [redacted] columns[70], each column containing many subfields. Google did not provide any explanation as to the content of the columns and what subfields are contained within. While I have attempted to describe in this report the many ways in which Google uses the private browsing information at issue in this lawsuit, this description is necessarily incomplete because Google has not provided full transparency in terms of its data collection and storage, and that information is not otherwise available.

173. A small subset of the columns in Google's DBL [redacted] match the names of logs stored in another Google data storage called [redacted] Assuming that these Google logs contain similar (if not the same) data, then Google's DBL [redacted] and Google's [redacted] store at least data based on user browsing activities that relate to user interest profiles for Google ad targeting as well as Google's user behavior profiles (such as a "clickiness" profile), which Google uses to count "how many times a user saw or clicked on ads".

[redacted]

[70] "2021-12-06 Brown - Special Master Submission.xlsx" tab "#6 DBL [redacted]"

174. Google provided some of Plaintiffs' DBL [redacted] data containing user profile information. These data were generated in Chrome Incognito mode while not signed-into Google. For example, [redacted] contains user profile keyed to Plaintiff Chasom Brown's Biscotti ID that Google generated from Brown's browsing activities on non-Google websites in Incognito mode while not signed-into any Google account.[71] This Google file (with Google collecting and then through the Special Master process providing [redacted] of data from Incognito browsing activities at the end of February and beginning of March 2022) contains a myriad of personal information and private browsing activities, including

[71] From Second Iterative Search of the Special Master process, "2022-03-14 Brown v. Google - DBL [redacted] – AEO" production for "Biscotti Extracted from IDE". The Biscotti ID value [redacted] is a hexadecimal value converted from Biscotti ID [redacted]. Decimal to Hexadecimal value conversion can be done using https://www.rapidtables.com/convert/number/decimal-to-hex.html for example (Last accessed on April 11, 2022). Additional support information can be found in Appendix H.1.

175. The browsing information that Google collected and then produced, just from DBL also includes keywords[72] on or related to websites that he has visited (e.g., at rows 116093-116796), affinity and in-market interest segments and a weighting factor for each segment, and more. Some of the in-market interest segments for Brown include among many others – the names of these segments are determined from GOOG-BRWN-00229632. Some of the affinity interest segments for Brown include , among many others – the names of these segments are determined from

[72] GOOG-CABR-04761793 at -796 explaining that Phil clusters are the "major concepts used on the page", Content verticals are the "primary topics of a page and the site", Felix criteria are "Keywords on the page or related to the page" and Label classifiers are "the tone of the page".

GOOG-BRWN-00229628. User profiles Google generated for other Plaintiffs while they were browsing in Incognito mode and not signed into Google contain similar information.[73]

176. Based on the productions made through the Special Master process and other discovery, Google does not appear to distinguish private and non-private browsing when it comes to user profile generation. The stored user profile information from Incognito sessions contains similar content as user profile information from non-Incognito sessions based on my review of Plaintiffs' data.[74] Indeed, Google explained in an Interrogatory response that, as to all of the data sources that Google identified during the Special Master process, Google's "uses [of the data stored therein] do not differ based on whether or not a user has enabled Chrome's Incognito mode" (Google's Response to Interrogatory No. 22).

177. On non-Google websites that use PPID, Google also stores user profiles keyed to a PPID-mapped-Biscotti ID. As discussed earlier, PPID and the corresponding PPID-mapped-Biscotti that Google maintains uniquely identifies a user on a particular non-Google website. The same identifier is used when a user signs into her non-Google website account on any device, in any browser and in both private and non-private browsing modes. Google explains that "The data collection and profile creation is namespaced by ppid & network_id" (GOOG-CABR-04122554

[73] See e.g., From the "2022-03-14 Brown v. Google - DBL [redacted] – AEO" production, [redacted] contains Plaintiff William Byatt's Incognito user profile information associated with Biscotti ID [redacted]. Additional support information can be found in Appendix H.2. As another example, [redacted] contains Plaintiff Jeremy Davis's Incognito user profile information associated with Biscotti ID [redacted]. Additional support information can be found in Appendix H.3.

[74] See e.g., From the "2022-03-14 Brown v. Google - DBL [redacted] – AEO" production, [redacted] contains Plaintiff Jeremy Davis's Chrome regular mode user profile information associated with Biscotti ID [redacted]. Additional support information can be found in Appendix H.4. As another example, [redacted] contains Plaintiff Monique Trujillo's Chrome regular mode user profile information associated with Biscotti ID [redacted]. Additional support information can be found in Appendix H.5.

at -571), where the "network_id" identifies the particular non-Google website. For example, [redacted] contains Plaintiff Monique Trujillo's user profile keyed to a PPID-mapped-Biscotti ID generated from Trujillo's browsing activities on [redacted] while not signed-into Google. Importantly, Trujillo logged into her [redacted] account while browsing in both Incognito and non-Incognito mode, including while signed out of Google, and all such browsing activity is used to generate her user profile associated with [redacted].[75]

178. Aside from DBL [redacted] other ads logs also contain user information, including sensitive information derived from visited websites in private browsing modes. For example, Google produced schema / fields for [redacted] ads logs, including signed-out display and search logs and signed-in display and search logs.[76] The schema contain fields such as: [redacted]. These sensitive fields are associated with a user's identity in GAIA signed-in logs and pseudonymous identity in signed-out logs.

[75] [redacted] is associated with PPID-mapped-Biscotti ID: [redacted], which is a hexadecimal value associated with PPID-mapped-Biscotti ID: [redacted]. In "2022-03-04 Brown Second Iterative Searches - PPID.xlsx", row 32, Google represented that this PPID-mapped-Biscotti ID is mapped from Trujillo's PPID [redacted] on [redacted]. As I show in Appendix H.6, Trujillo logged into [redacted] in Incognito and non-Incognito mode while not signed-into Google and the same PPID is sent to Google in both modes.

[76] 2022-03-11 Production containing schema from searching [redacted] logs using three Plaintiff IDs for each log.

E.2 Advertising: Throughout the class period, Google used private browsing information to serve ads (including targeted ads)

179. Throughout the class period, Google used the private browsing information at issue in this lawsuit to serve users with ads (including targeted ads) in that private browsing session (and thus earn money for Google from serving those ads). This was demonstrated by the various tests performed at my direction in connection with the Special Master process. Further, as Google's internal document explains, "Targeted advertising occurs based on data from current Incognito session (tracking) cookies can be placed while Incognito" (GOOG-BRWN-00147873 at -877).

180. Targeted ads can be personalized or non-personalized and can appear on non-Google websites, for example as banner ads or within embedded YouTube videos. "Google considers ads to be personalized when they are based on previously collected or historical data to determine or influence ad selection, including a user's previous search queries, activity, visits to sites or apps, demographic information, or location. Specifically, this would include, for example: demographic targeting, interest category targeting, remarketing, targeting Customer Match lists, and targeting audience lists uploaded in Display & Video 360 or Campaign Manager 360… Non-personalized ads are ads that are not based on a user's past behavior. They are targeted using contextual information, including coarse (such as city-level, but not ZIP/postal code) geo-targeting based on current location, and content on the current site or app or current query terms" (GOOG-BRWN-00016645).

181. Google's internal documents confirm that Google personalizes advertisements based on private browsing activities, with Google employees recognizing (in connection with Chrome) that "Users overestimate the protections that Incognito provides and are unaware of the personalization and data collection that occurs when it is on" (GOOG-BRWN-00529122). As another example,

Google employee AbdelKarim Mardini confirmed that, "during the time-span of one incognito session personalization can happen based on the activities happening in that incognito session" (Mardini November 24, 2021 Tr. [294:4-8]). Google likewise admits that when Google receives data from a user's Incognito session, "Google may use the data associated with cookies to personalize advertisements displayed to the user." (Google's Response to Request for Admission No. 7).[77] And when asked during his deposition "How do cookies within Incognito mode help monetize a publisher's site," now-retired engineer Justin Schuh answered "the same way they do when you're not in Incognito mode. They give you or they – they can be used to deliver targeted advertising' (Schuh January 6, 2022 Tr. 156:14-24).

182. Google personalization, in terms of targeted advertisements, includes advertisements by any advertiser "in any private browsing mode on any browser." (Mardini November 24, 2021, Tr. at [294:9-13]). Google's internal documents also reveal Google's concerns about such personalization of private browsing, noting "privacy risks: … Session-based personalization, Use of non-personal data for personalization, i.e. location" (GOOG-CABR-00411167 at -170).

183. To implement personalization, Google uses [redacted] targeting label sources, [redacted] targeting machine learning models, and [redacted] data pipelines to generate audience lists for personalized advertising. These labels include past user interests (e.g., past ad clicks, conversions), current user interests as well as predicted future interests (GOOG-CABR-04616196 at -215, -217, and -222). Google's machine learning algorithms use various signals, including contextual signals, location, browser, etc. to predict users' future interests (GOOG-CABR-04616196 at -221 to -223). Predicted future interests account for more than half of Google's Display Ads revenue and are used to

[77] *See also* Google's Response to RFA No. 11 ("Google admits that . . . Google may have been able to display ads using data Google received while a user was in Incognito mode."); Google's Response to RFA No. 40 (similar).

compensate for cookie loss (GOOG-CABR-04616196 at -223). Machine learning tools relevant for ad serving include Sibyl for ad ranking, Laser for personalization, and SmartASS[78] (Smart Ad Selection System). Google tells its new employees that SmartASS "Makes a LOT of money. This most likely pays your salary!" (GOOG-BRWN-00535100 at -156). There are many SmartASS related fields in the ads log schema produced through the Special Master process.[79] The above machine learning systems utilize data that Google collects about users' online activities, including users' private browsing activities.

184. When a user visits a non-Google website with embedded Google advertising code, ad request messages are sent to Google along with the user's identifiers for that browsing session. However, Google does not just serve ads based on those identifiers alone. Its ad serving infrastructure takes one or more IDs from the current session and "retrieves all associated IDs" (GOOG-CABR-04126517 at -522). Such ID associations (achieved through linking identifiers Google estimates to belong to the same user on the same device and across different devices) are "used extensively in ads personalization" and other ads related functions (GOOG-CABR-00543864 at -884 to -890 and -902 to -905). Through Google's server-side processes such as the [redacted][80]), Unified ID linkage Storage (for storing and querying ID linkage data[81]),

[78] Google seems to fancy ironic acronyms.

[79] For example, the following fields are in the log schema of the logs containing the "maybe_chrome_incognito" field: [redacted] (2022-03-10 Brown v. Google – Fields for [redacted] Logs – AEO.xlsx)

[80] GOOG-CABR-00543864 at -872 and -877

[81] GOOG-CABR-00399988 at -991 and -992

[redacted][82]) and others, Google creates "a single-view of the user in all Display products" for ad targeting and measurement purposes (GOOG-CABR-03652751 at -752).

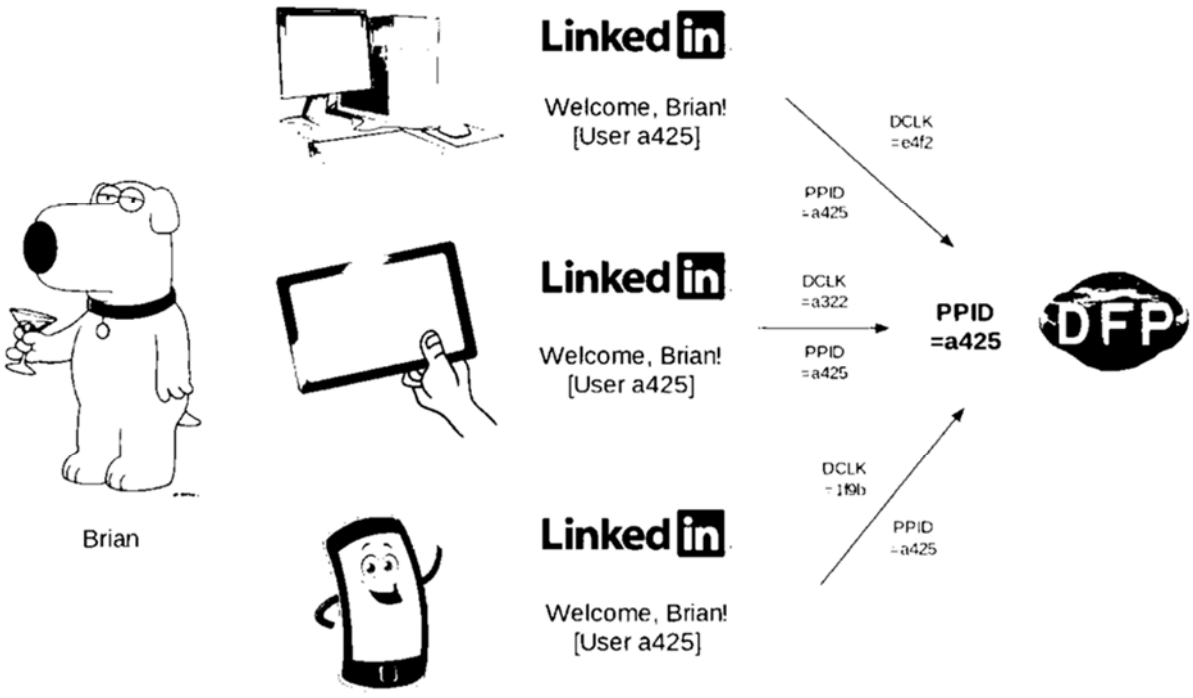
185. As a part of serving targeted ads, Google uses its dynamic remarketing tracking beacon on non-Google advertiser websites to send Google GAIA or Biscotti cookies along with a "list of products on the page (as keyvalues)" (GOOG-CABR-04763333 at -336). Google then stores this data (from the user's browsing activities) in Google's [redacted] data repositories. "Later, when the user is on a publisher page, ad request with the [GAIA or Biscotti cookie] is sent to Google server. If a dynamic remarketing adgroup wins, then [Google] will fill the creative with products user saw and other 'related products' based on [machine learning] models" (GOOG-CABR-04763333 at -337).

186. While dynamic remarketing aims to show products to users who have visited an advertiser's website, Google uses what it calls "dynamic prospecting" to generate Google revenues by showing products to new users. "Dynamic prospecting uses machine learning to get an idea for what potential buyers are looking for. Once the system knows what the user is after, it combines that possible intent with demographics-based information such as age, gender, and household income to match the user with a product" in an advertiser's list of products and services uploaded to Google.[83]

[82] GOOG-CABR-03652751 at -752: [redacted], -756: [redacted] and at -769 "Infer Same Device Links from [redacted] – Build heuristic rules based on cookie attributes – Achieve high prediction accuracy".

[83] GOOG-CABR-05876730 and GOOG-CABR-05876831

187. With the loss of certain Google cookies (e.g., due to blocking of certain Google cookies based on Apple's Intelligent Tracking Prevention ("ITP") or Google's [redacted] changes to Chrome Incognito mode), Google mitigates targeting loss by relying on first-party identifiers, such as the PPID available for Google Ad Manager 360 Publishers. This identifier is used to identify users that log into publisher websites, and it uniquely identifies a user across all of the user's devices, browsers, and browsing sessions, including private browsing sessions. While for purposes of Chrome Incognito, cookies such as Biscotti are specific to a particular session, PPIDs, on the other hand, are similar to GAIA in that it is a persistent identifier for a particular user used across sessions and devices. This is illustrated in the screenshot below from one of the documents produced by Google in this litigation, showing different DCLK (Biscotti IDs) but the same PPID across different devices (GOOG-CABR-04609856 at -857) for the same publisher (in this case, LinkedIn).



188. PPIDs can be set by the publisher or managed by Google ("Managed PPID"). Managed PPIDs are dropped by Google's tracking beacons through JavaScript (GOOG-CABR-04118321 at -333). PPIDs are used to "create user profiles from traversal history of the user on sites in the publisher network" and for remarketing ads (GOOG-CABR-04118321 at -330). The "traversal history of the user on sites in the publisher network (can be cross device, and cross domain)" (GOOG-CABR-04118321 at -331).

189. Google explains that the "main benefit of using PPID as the basis for an Audience segment member is to be able to target users across devices and browsers, or in cookie-less and similar environments… The PPID is mapped to a new biscotti ID the first time [Google sees] an ad request where the publisher specified a PPID. The PPID is hashed and stored in a table mapping PPIDs to biscotti IDs. Any time a PPID is specified in the request, the ad selection will consider the audience segments associated with the mapped biscotti ID and DoubleClick cookie if the user also has a separate DoubleClick cookie on the browser" (GOOG-CABR-04609856 at –859 and at -860); *see also* Berntson March 18, 2022 Tr. 131:4-9 ("[S]o if the same publisher is interacting with you on the web and interacting with you on your phone, we'll see you as the same person across those devices for that one publisher with PPID. And we designed it that way.").[84]

190. The implication of PPIDs for the private browsing at issue in this litigation is that the same PPID-mapped-biscotti is used by Google to key all user data associated with a user's private and non-private browsing sessions on a publisher website. Thus, a user's browsing activity within a

[84] *See also* Berntson March 18, 2022 Tr. 120:17-22 (describing the PPID as something that "is, to the publisher, a unique representation of the user"); id. 126:4-6 ("[O]ne of the things that is used fairly commonly ***when there is a way to identify a user, like with a PPID***, is frequency capping . . ." (emphasis added)).

private browsing session can be used by Google to create an even more robust user profile and personalize ads in a non-private browsing session.

191. Aside from ads targeting, Google also uses (and throughout the class period has used) the private browsing information at issue for ad frequency capping (i.e., limiting the frequency of showing a particular ad), sequencing advertisements and spam detection (GOOG-CABR-04423260 at -262, -277). PPID, and another 1P Google ID set by the Ad Manager tracking beacon in a GFP cookie, are both used for frequency capping.[85]

E.3 Conversion Tracking: Throughout the Class Period, Google used private browsing information to track conversions

192. Throughout the class period, Google also used the private browsing information at issue in this lawsuit to track what Google refers to as "conversions." A conversion is any action, or goal, that an advertiser seeks to encourage through online advertising. A conversion "occurs whenever a user takes any action that is identified as important or relevant to a given advertiser/marketer," (GOOG-CABR-03738741 at -41) including signing up for something, making a purchase, or filling out a web form.[86] Google's conversion tracking applies alike to ads shown on publisher websites (display ads conversions), YouTube (YouTube ad conversion), and Google Search (search ad conversion). *See* Berntson March 18, 2022, Tr. 154:22-155:4 (noting that the process for tracking display conversions relative to Google Search conversions is, "if it is not the same, it is similar").

[85] GOOG-CABR-04118321 at -330, -337, -339 and GOOG-BRWN-00162235 at -242.

[86] *See also* March, 18 2022 Berntson 30(b)(6) Tr. 62:2-8 ("conversion generally, in the advertising space, refers to if an ad is displayed that wants the user to take some action, did the action take place? The vast majority of cases, what ads are doing in terms of seeking that conversion, is to purchase something").

193. Conversion tracking has been and continues to be one of the primary ways for Google to demonstrate its value to advertisers (meaning, to justify the payments Google receives from advertisers). Google's conversion tracking is also heavily reliant on Google obtaining and analyzing user browsing information, including from users' private communications with non-Google websites in private browsing modes.

194. Google's conversion tracking focuses on tracking user activities. For example, if a user is shown an advertisement and then makes a purchase, that could qualify as a conversion. Another example of a conversion might be a user clicking an ad, arriving at the advertiser's website, and then signing up to receive special offers via email. Advertisers are generally willing to pay more for advertising when Google can demonstrate the effectiveness of its advertising by tracking conversions (Kleber January 14, 2022 Tr. 168:1-19). Appendix F contains a discussion of conversion tracking history at Google and conversion tracking methods.

195. Google's conversion tracking involves tracking user activities on non-Google websites. When a user sees an ad, clicks on an ad or otherwise interacts with an ad (e.g., watches a video ad) on non-Google or Google websites alike and later performs a conversion action on an advertiser's webpage or offline, that user is considered to have "converted". User activities on non-Google websites are tracked by Google tracking beacons embedded on advertiser webpages[87], including users' private browsing activities. For example, "The user clicks an ad, they are redirected to a Google-owned domain where a cookie is set, then finally redirected to the customer's landing page [a non-Google website]. When a user later converts [i.e., buys something on the website], the conversion tracking beacon picks up the conversion cookie from the Google-owned domain (which in that context is third-party [i.e., non-Google]) and the conversion can then be joined back

[87] GOOG-CABR-05876763

to the original click." (GOOG-CABR-00903677 at -77). Google tracks conversions by way of Google conversion tracking code embedded within non-Google websites, which cause the user's browser to transmit to Google the information contained within the HTTP request, including the user agent (Berntson March 18, 2022 Tr. 76:23-77:2). Appendix A includes a discussion of Google's conversion tracking code.

196. Tracking conversions is an important part of how Google makes money, with Google using that information to prove to the advertiser that Google's advertising campaign is working and that it therefore makes sense for the advertiser to pay Google for the service. For example, consistent with my analysis and opinions, one Google employee testified:

> Showing an advertiser that a conversion happened is a way for the advertiser to realize that this particular ad was a good ad for them to have chosen to show, right? So advertisers have to make decisions about how to spend their money, and I think they tend to spend their money on advertising that seems to be a good financial investment for them, right, where the money they make later on is worth the advertising cost that they invested. So knowing that a conversion happened is important to advertisers because it lets them better choose how to allocate their ad spending budget in the future.

Kleber January 14, 2022 Tr. 168:5-19.

Similarly, testifying as Google's corporate representative, Dr. Glenn Berntson explained:

> And specifically, Google collects the conversion data because when an advertiser comes to Google says, "You know, I -- I want to reach N number of users and I want to drive sales for this particular product. So Google, I want to use your system to serve an ad," whether that's Ad Words or DV3, in order to for Google to actually meet the objectives of the advertiser, we have to be able to measure the conversions that map to the objectives of the advertiser, and then be able to provide that data back to the advertiser.

Berntson March 18, 2022 Tr. 63:12-23,

197. Dr. Berntson likewise testified during a prior 30(b)(6) deposition that "an advertiser will create a campaign that says, I'd like to serve this 7-up ad and here - here are the criteria for where I want the ad to show up . . . and then if they want to actually see how many people then buy that soda because they saw the ad, what we provide back to the advertiser is not a list of IDs of users

that converted, but just how many conversions happened because of their campaign." (Berntson June 16, 2021 Tr. 200:6-19). Google's ability to highlight these conversions for its advertising customers depends on Google's systematic tracking of users' browsing activities, including the private browsing activities at issue in this lawsuit over the class period.

198. When Google provides conversion statistics to advertisers for purposes of "being able to demonstrate to an advertiser with we helped them meet their objectives", Google "basically just share[s] an aggregate number" (Berntson March 18, 2022 Tr. 64:3-21).

199. Google's conversion tracking continued throughout the class period, with limited changes. In 2017, Google unified Analytics and other conversion tracking beacons (e.g. AdWords conversion tracking and Floodlight) under the One Google Tag (OGT), also called Global Site Tag (gtag.js) – thus "gtag.js is Google's new conversion tracking, analytics, remarketing etc. tag – all in one".[88] Google tells websites to install this Google tracking beacon on every page of their website (a process known as site-wide tagging).[89] Site-wide tagging can also be implemented through Google Tag Manager (GTM) and Google Analytics. Google explains the benefits of site-wide tagging as providing "more accurate conversion tracking" (GOOG-BRWN-00026913 at -920). "When 3P [e.g., Google cookies that are blocked by third-party cookie blocking browsers] cookies are not available, sitewide tags allow accurate conversion measurement with 1P [e.g., Google cookies that are not blocked by third-party cookie blocking] cookies" (GOOG-BRWN-

[88] GOOG-CABR-04419756 at -761 and at -756: "gtag.js aims to unify Google's ads & marketing site measurement tools under a single implementation, where advertisers can use a common tag on their sites to implement various analytics and advertising product features, such as conversion tracking and remarketing. Currently, different Google products use different tags. This has led to a confusing, fragmented ecosystem of Google products for our customers. gtag.js consolidates all of these differences into a single tagging solution, making it significantly easier to measure website actions and send that data to the correct location."

[89] https://developers.google.com/analytics/devguides/collection/gtagjs (Last accessed on April 11, 2022).

00026913 at -921).[90] Appendix D discusses the difference between 1P and 3P cookies in more detail.

200. Google's tracking code, together with cookies and parameters that Google embeds in URLs, can be used by Google to connect conversion events on non-Google websites with particular ads seen or clicked by a user, regardless of whether that ad was seen or clicked on Google's search results page, in YouTube videos, or on other non-Google websites. GOOG-CABR-04324934 describes tracking conversions for Search ads, YouTube ads and Display ads using sitewide tagging, conversion cookies, and server-side conversion tracking (SSCT).

201. Google's conversion tracking (including with the private browsing information at issue in this lawsuit) plays an important role in generating Google revenue. As noted above, Google uses conversion tracking to justify advertising spend. Advertising generates [redacted] of dollars in revenue for Google every year. As one example of the value Google recognizes from conversion tracking, focusing on sitewide tagging from Google Analytics alone in 2018, those "conversions are used in powering [redacted] [Google Ads] auto-bidding spend" (GOOG-CABR-04081967).

202. Moreover, in the process of tracking conversions, Google amasses valuable data to feed Google's machine learning algorithms, the engine to Google's [redacted] advertisement business. Since machine learning algorithms build models by learning and generalizing from data. Even if data is not used to monetize transactions related to a specific user, the data helps Google better monetize the transactions of its user population, or of users similar to the specific user.

[90] *See* Chung Tr. 69: 3-12; 16-18 (describing addition of gclid to the GA cookie which can be used by Google advertising products "along with the data coming in with Analytics—to parse and determine whether a conversion occurred" even when third-party cookies are blocked); *see also* Chung Tr. 76:14-77:14 (confirming that even when third-party cookies are blocked in Incognito mode, Google Analytics still receives and stores in logs "the interaction, the behavior that the visitor executes on the website").

203. Conversion data can also be monetized through manual or auto bidding. Manual bidding requires the "Advertisers download conversion reports and use it to manually adjust bids or budgets" (GOOG-BRWN-00144320 at -330). Advertisers may also view conversion data in the Google Ads online interface and then manually adjust their bids. In my opinion based on my analysis in this case and my experience as a Google Ads advertiser, Google does not distinguish between private and non-private browsing activities in those Google Ads online interfaces, with Google providing customers with comingled conversion data without separately identifying private browsing conversions. I have not seen or heard of an option to segment the data reported by Google to show private browsing and standard browsing activity separately, despite the fact that Google offers a very large number of different ways to segment data in both Google Ads and Google Analytics reports.

204. Bids can also be adjusted automatically through machine learning algorithms.[91] With autobidding, "advertisers describe the outcomes that they want, rather than being in full control of the targeting, and then say to Google, 'You figure out how to target this campaign. You figure out the best places to show ads, and I want you to maximize my outcomes,' which could be up -- you know, the number of conversions" (Berntson March 18, 2015 Tr. 69:11-18).

205. Google's machine learning algorithms analyze millions of user signals, including users' browsing history, search queries, location, device, age, gender, demographics, interests, operating system, browser and much more, all "within the blink of an eye" (GOOG-CABR-04786706 at -732). In such a machine learning system, knowing the characteristics of a user who does not

[91] Machine learning algorithms are used across a wide array of Google products, including generating user profiling (e.g., finding similar audiences, in-market audiences, detailed demographics, affinity audiences), creating responsive display and search ads that tailor the ad message "for every moment, for every user", and "adjusting bids for each and every [ad] auction" (GOOG-CABR-04786706 at -729 and -730)

convert is valuable in predicting which users will convert, and thus allows Google to maximize the value of its advertising program for itself, its advertisers and its publishers. As of year 2020, Google was the leader in online advertising in the United States with a 28.9% market share.[92]

206. One type of conversion is a same-session conversion, which can occur while a user is in Chrome Incognito mode. *See* GOOG-CABR-00547295 at -95 (referring to a "conversion [that] happens in the same incognito session"). A same-session conversion occurs when a user views an advertisement and then takes some desired action within the same browsing session, including the same Chrome Incognito browsing session. Importantly, the user need not convert right away after viewing or clicking on the ad. For example, within a single Chrome Incognito browsing session, a user may browse on a non-Google website, see an ad, later see the same ad on a different non-Google website, click on it and make a purchase. A user may also see an ad on a non-Google website, decide to do some research on the product by doing some searches and additional browsing, before eventually making a purchase. *See* Rakowski Tr. 191: 13-19 ("If it starts in Incognito and completed the conversion in Incognito, that would appear normally in the Adwords conversion track…tracking report.").

207. Google has implemented various mechanisms to track these conversion events, including in Chrome Incognito mode. Within the same Chrome Incognito browsing session, Google's conversion tracking works the same way as conversion tracking in non-private Chrome browsing mode.[93] Google further explains, "On the path to conversion, customers may interact with multiple

[92] https://www.statista.com/statistics/242549/digital-ad-market-share-of-major-ad-selling-companies-in-the-us-by-revenue/ (last accessed on February 17, 2022)

[93] GOOG-CABR-03662975 at -978: "I was reading the latest on Chrome blocking 3P cookies by default in incognito mode and its effect on measurement. … I understood that we were able to track same session conversions in Chrome incognito mode across 1P/3P cookies"; *see also* Berntson March 18, 2022 Tr. 145:7-14 ("I don't think it's possible for Google Scripts -- whether it's Google scripts supporting Ad

ads from the same advertiser… Most advertisers are used to measuring the success of their online advertising on a 'last click' basis. This means they give all the credit for a conversion to the last-clicked ad".[94] Other attribution models are also possible, such as equal attribution for all clicked ads, or all credit to the first clicked ad.

208. Google also tracks cross-device or cross-session conversions. These conversions occur when a user completes the conversion event on a different device (or within a different private browsing session) from the device/session in which the user initially viewed the advertisement. For example, within a private browsing session, a user may browse on a non-Google website on a laptop, see an ad, later browse on a mobile device and make a purchase related to the ad she viewed earlier. A user may also see an ad while browsing in one private browsing session and later convert in a different private browsing session. In these cases, Google tracks the conversion using persistent identifiers or personally identifiable information (PII) across different devices or sessions.

209. For example, through Enhanced Conversions in Analytics and Ads, non-Google websites send personally identifying information such as a user's email address (user's name, home address, and phone number may also be used) to Google to be matched against the same user's Google account information containing the same identifying information along with the user's GAIA ID

Manager, Google Analytics, or our advertiser products, OGT -- I don't think it's possible for those client-side scripts to change what they do based on the knowledge of whether private browsing has been enabled in Chrome").

[94] https://support.google.com/google-ads/answer/6259715?hl=en (Last accessed on April 11, 2022) "Note: Currently, attribution models (both "Data-driven" and non-last click models) are available for Search campaign performance reporting and bidding optimization. In the coming months, we'll be upgrading attribution models to also cover ad interactions on YouTube and Google Display Ads. As a result, you might see conversion credit shift between and within your Search, Shopping, YouTube, and Google Display Ads campaigns. Some traffic will still be reported using the "Last click" model, such as Display campaigns using pay for conversions billing."

(GOOG-CABR-04763527). With Enhanced Conversions, Google can track cross-device conversions when a user performs a conversion event in private browsing mode and is not signed into Google.

210. Even when Google is unable to track all of the conversions, Google recovers these "lost conversions" through conversion modeling.[95] While "measur[ing]" conversions refers to "actually see[ing] that [a conversion] happened and account[ing] for the fact that a conversion happened," "Modeled conversions are conversions that we estimate took place without actually measuring them." (Berntson March 18, 2022 Tr. 65:4-16). For example, "imagine somebody walks into your store and you see how tall they are, and then you can go look up this historical data and see what the relationship is between height and arm length, 'Oh, they're that tall, I bet their arm is this long.'" (*Id.* 66:18-23). "Sometimes, [Google] will be able to . . . make those measurements of the conversions and build such a model to infer the likelihood of a conversion having taken place without being able to measure it directly" (*Id.* 67:11-16).

211. In a September 2020 email, a Google employee explained that Google was developing a "modeling gap . . . to fill all the conversions that do not happen within the same incognito session" (GOOG-CABR-00547295 at -295). One tool that Google uses is Google's [redacted] program.[96] With [redacted]

[95] GOOG-CABR-00547295 at -295: "Chrome in incognito still supports 1P cookies, without TTL. So, the modeling gap is to fill all the conversions that do not happen within the same incognito session".

[96] GOOG-CABR-04959606 at 609: "Incognito: Users in incognito mode usually don't sign-in with gaia, thus their clicks/conversions are all sign-out. Those sign-out XD [cross-device] convs are already recovered by [redacted] XD modeling."

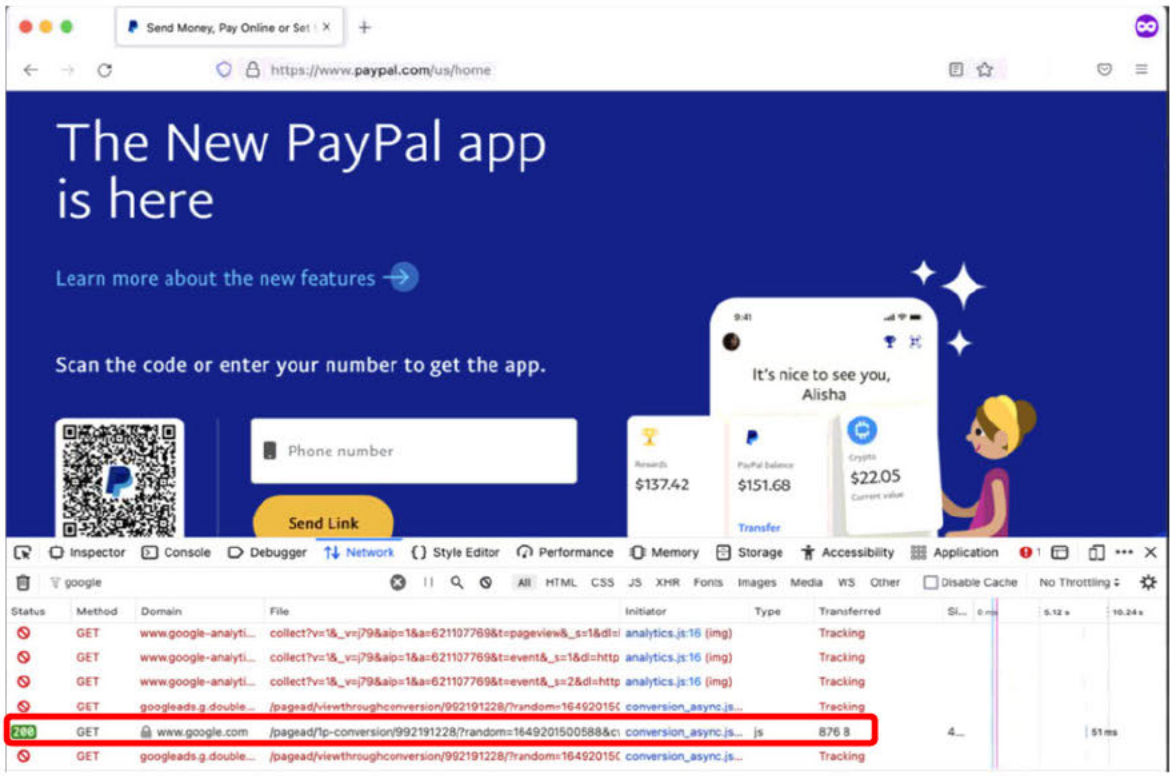
(GOOG-CABR-03670580). See Appendix F for a discussion of [redacted] as well as other cross-device conversion modeling tools used by Google.

E.4 Google has attempted to circumvent the blocking of Google tracking beacons by Firefox's private browsing mode and blocking of cookies by Apple's Intelligent Tracking Prevention (ITP)

212. As discussed above in Section A, Firefox private browsing mode attempts to block, by default, all third-party trackers, including Google tracking beacons. Rather than praise Firefox for actually attempting to provide users with a private browsing experience, or adopting similar functionality for Chrome Incognito, Google has viewed Firefox private browsing as a problem—focusing on how its features diminish Google revenues. For example, in a July 2019 internal Google presentation titled "Firefox Private Browsing", Google employees discussed a "conversion tracking gap in FF private browsing" and the estimated impact of "[redacted] of Google's revenue" "once google.com gets blocked by FF" (GOOG-CABR-04695140 at -143, -148 and -151). The Google presentation specifically discussed ideas for a "short term client side fix" as well ideas to implement "ground-truth for modeling" (Id. at -143, -151). With this presentation, Google employees acknowledged that Google was still generating revenues from its collection of Firefox private browsing sessions, and Google's desire to continue generating those revenues even once Firefox made changes to no longer permit Google to collect that information.

213. The work around Google devised to circumvent the protections afforded to users with Firefox private browsing mode is to send conversion tracking information to a google.com domain that is not blocked by Firefox (GOOG-BRWN-00027314 at -320). For example, I have observed conversion tracking information being sent to Google even in Firefox private browsing mode. As

shown in the screenshot below, visiting paypal.com while in Firefox private browsing mode, conversion information sent to google.com is not blocked (this is highlighted in the red box below).



214. Google's circumvention of the Firefox private browsing protections is not an isolated response by Google. Rather, Google has consistently responded to other browser's privacy features by developing work arounds. For example, in response to Apple's Intelligent Tracking Prevention ("ITP") initiative, limiting tracking via cookies, Google developed and has used several work arounds under an internal project called [redacted] (GOOG-CABR-04697357). As Google states in GOOG-CABR-05362596 at -596: [redacted] I discuss these Google work arounds in Appendix F.

E.5 Analytics: Throughout the class period, Google used private browsing information to provide information to Google's Analytics customers and increase Google's total ads revenues

215. In addition to ads (discussed above), Google throughout the class period systematically tracked user activities and benefited from that tracking by way of Google Analytics, including with regards to the private browsing information at issue in this lawsuit. "Google Analytics is the most widely used analytics solution in the world…covering about [redacted] of the web" and it is "one of the largest products at Google" (GOOG-BRWN-00424253 at -279 and at -295). As of 2021, Google maintained [redacted] of the total web analytics market share.[97] As shown in the figure below, as well as in internal documents throughout the class period, Google Analytics is installed on [redacted] of the top [redacted] websites and is a significant contributor to Google's [redacted] advertising business.[98] Within the United States, "[t]here are [redacted] [Google Analytics] Accounts," and "[a]ccounts may be associated with more than one website" (Google's March 17 Response to Interrogatory No. 33).

216. There are two versions of Google Analytics, one offered for free and one where customers pay Google (Google Analytics 360), with annual revenue from Google Analytics 360 totaling over [redacted] in 2020 and Analytics impacting [redacted] of dollars of Google's advertising

[97] https://www.statista.com/statistics/1258557/web-analytics-market-share-technology-worldwide/ (Last accessed on April 11, 2022); *see also* Chung Tr. 139:1-5 ("That [redacted] is the market penetration I've been told over the past year and a half I've been on the Google analytics team. . . To my understanding, this would still be true.").

[98] GOOG-BRWN-00490767 at -772 (Last updated Nov. 2020). Similar statistics are shown in GOOG-CABR-00381312 at -321 (Last updated Nov. 2019), GOOG-BRWN-00491711 at -718 (Last updated Mar. 2018), and GOOG-BRWN-00424253 at -279 and -295 (Last updated Sept. 2017).

business. Google's internal presentations confirm the broad adoption of Google Analytics and benefit to Google, including:

217. Google Analytics' key role in generating advertising revenue for Google is reflected in other internal documents. For example, on internal Google presentation notes: "Google Analytics is a free service, but that doesn't stop it from making money. Rather than providing a direct source of revenue (by charging for the service), Analytics generates more Google ad revenue. It's pretty simple: the more successful an advertiser considers an online ad, the more money that advertiser is willing to spend on that ad and similar ads. Google Analytics offers the information that can prove the success of those ads" (GOOG-BRWN-00026812 at -813). Providing advertisers proof of efficacy is a well-established method for an advertising network, such as Google, to increase revenues.

218. As discussed elsewhere in this report, Google has tightly integrated Google Analytics with Google's ad serving infrastructure.[99] Google designed its systems so that, throughout the class period, analytics data has been shared with ads in real time through a series of HTTP redirects (i.e., join beacons) shown below for targeted advertising and remarketing (GOOG-CABR-05404845 at -855).[100] Google refers to these join beacons as [redacted] and [redacted] (GOOG-CABR-05404845 at -863).

[redacted]

[99] GOOG-CABR-04678169 (Last updated: 2013/06/19) at -169 describes Analytics and ads (DoubleClick) integration through a "dual-beacon system" and at -170: "By including the DoubleClick integration in analytics.js directly, we risk 'tainting' Google Analytics, which has a reputation for anonymous web analytics tracking via first-party cookies, with DoubleClick, which does not.

[100] *See* Chung Tr. 111: 8-9, 21-22, 112: 13-18 (affirming Analytics client ID exists in a [redacted] with "DoubleClick or Biscotti" advertising identifiers and identifying at least three such [redacted] from GOOG-BRWN-00029378)

219. Remarketing is an important part of Google's business, and how Google makes money, providing a way to serve targeted ads to users who have visited particular non-Google websites. There are two types of remarketing:

- Remarketing Lists for Search Ads (RLSA) where search ads are shown on google.com to past visitors of non-Google websites. For example, suppose a user visits a "website that sells safety razors, www.wilkinsonsword.co.uk, but didn't buy. The next time they look up in Google, say, 'blades', they're shown" Wilkinson ads (GOOG-CABR-00377968).
- Standard Google Display Network (GDN) Remarketing where remarketing ads are shown on other non-Google websites. For example, the same user, having visited www.wilkinsonsword.co.uk now visits "beardedgospelmen.net, and there they're shown [Wilkinson] ads about … safety razors – as [Google] assume that a website talking about beards is a place where people may be interested in seeing these type of ads" (GOOG-CABR-00377968).

220. Google's Analytics and advertisement tracking beacons both track users to create "Remarketing lists" for use by Google. Users are added to particular remarketing lists based on their past browsing activities on non-Google websites, including browsing in private browsing modes. In the case of Remarketing List for Search Ads (RLSA), a user's Zwieback cookie is added to the list and in the case of Google Display network (GDN) remarketing, a user's Biscotti cookie is added to the list (GOOG-CABR-00377968 at -972). These lists are used during ad serving time to serve remarketing ads to the user.

E.6 Google also used private browsing information to enhance Search revenue and develop and improve other Google products and algorithms

221. Google's internal document indicates that "users exposed to display ads might search more" (GOOG-CABR-04618590 at -593). Thus, users browsing activities on non-Google websites not only generate Display ad revenue for Google but also generate more Search ad revenue as well. In addition, as discussed earlier, remarketing is a powerful tool to increase search ad revenue based on users' past visits to non-Google websites.

222. Google likewise admitted that "Google has used some of the data it collects, including some of the data at issue in this case, to improve its services." (Google's Response to Request for Admission No. 47). I am not aware of Google providing any list of all of the ways Google benefits, such as all of the "services" that have purportedly been improved, but one key way in which Google develops and improves its various services is with data-driven algorithms and development.

F. Google Joinability: Throughout the class period, Google collected and stored private browsing information in ways that can be joined to other Google user information, but Google withheld and destroyed data that would be relevant to further assessing and demonstrating that joinability

223. Based on my analysis of the case discovery, including with the Special Master process, and my experiments, it is my opinion that Google throughout the class period and across the two classes systematically collected and stored detailed private browsing data that constitutes what is commonly understood to be sensitive fingerprinting data, which can be used to identify users and join data. As Google explains, "Fingerprinting is worse than cookies in some aspects: (1) fingerprinting normally doesn't store anything onto your computer, so you can't clear or reset the tracking data, (2) fingerprinting is done silently without users' consent" (GOOG-CABR-04006287 at -288). Regardless of whether or not Google actively "fingerprints" its users, the voluminous amount of fingerprinting data it collects and stores – including from private browsing – certainly allows data to be tied to individual users.

224. Some of the fingerprinting data collected by Google from private browsing throughout the class period include IP address, user agent, type of browser, device, screen resolution, location, time zone, language, cookies and more.[101] My opinion that this constitutes fingerprinting data is

[101] For example, from the Second Iterative Search, production 2022-03-14 Brown v. Google - [redacted] – AEO, file [redacted] contains Plaintiff William Byatt's Incognito

consistent with the sworn testimony by several Google employees as well as many internal Google documents discussing private browsing and fingerprinting, as discussed below.

225. For example, one Google internal document confirms that information obtained from Incognito browsing can be used for fingerprinting, noting "Incognito mode and regular mode should not be linkable…This is to some extent violated by fingerprinting" (GOOG-BRWN-00047390 at -392). Further, Google's internal documents acknowledge that "The idea that users behind Zwieback UID is not identifiable if one were to use all of the data we have available is laughable." (GOOG-BRWN-00433503 at -503). Discussing this document during his Rule 30(b)(6) deposition, Google engineer Michael Kleber explained that "from the point of view of Zweiback [Zwieback] cookies, . . . they work the same in Chrome Incognito mode and Chrome regular mode and other browsers. It's all the same" (Kleber March 18, 2022 Tr. 66:16-23).

226. Another internal Google document explains that "data collected while Incognito is strictly **NOT** anonymous. When using a Google service, data (e.g. search queries) is tied to a pseudonymous ID (e.g. zwieback) that is then stored with the user's IP address, user agent, and

---

browsing data based on a search of his Biscotti ID: [redacted] within the [redacted]. Each event record includes the time the event is generated, the URL of the page within the [redacted] field, IP address within the [redacted] field, "Referer:", "UserAgent:", language within [redacted] and [redacted] fields indicating the language is [redacted]. Biscotti cookie value that Google used to find this log entry, and other information, such as ad query and ad click information. As another example, an Analytics log called [redacted] contains Plaintiff William Byatt's Incognito browsing data based on a search of his Analytics CID: [redacted]. This log, from production 2022-03-15 Brown v. Google - [redacted] – AEO, contains visited [redacted] among many other pieces of information. Analytics CID [redacted] and Biscotti ID [redacted] are shown to be associated with the same event in production 2022-03-25 Brown v. Google – Analytics [redacted] data – AEO, file [redacted] Row 88. Column M of Row 88 contains the CID and Column A of the same row contains Biscotti cookie value: [redacted]. Google shows that the Biscotti ID embedded in this cookie is [redacted] in "2022-03-02 Brown v. Google - Decode IDE.pdf", page 1, item 9. Additional support information can be found in Appendix H.7.

other metadata (everything in the gwslog proto). While we don't connect it to a user's identity and there are many internal policies to ensure it does not happen, it's not impossible. Further, under the GDPR [General Data Protection Regulation], data tied to unauthenticated identifiers **IS** considered to be personal data… That means everything a user does while signed out of Google services is considered to be personal data" (GOOG-CABR-00501220 (emphasis in original)).[102]

227. IP address and User Agent are important fingerprinting data, which can be used to identify users. As Google explains, "IP address is a viable tracking vector and as such poses a risk to the privacy of users browsing the web." (GOOG-BRWN-0072761). Google's internal documents acknowledge that "IP addresses often uniquely identify users" (GOOG-BRWN-00613409), "IP address + UA (user agent) can reveal individual user with high probability." (GOOG-CABR-03611484 at -485), "the User-Agent and IP address will often uniquely identify a particular user's browser" (GOOG-CABR-04006287 at -288), and "'combining IP with other data' can be trivial when it comes to things like your user agent" (GOOG-BRWN-00157001) and "connect regular and incognito sessions…doable using IP addresses and if Google turns evil" (GOOG-CABR-05737898 at -901). These documents are consistent with my analysis and opinion that Google can readily use the data at issue in this lawsuit (collected from private browsing during the class period) to identify users and their devices.

[102] *See* GOOG-BRWN-00572220 at -248 (noting that IP addresses "can be used as join keys for linking Incognito:Zwieback activity to GAIA"); GOOG-CABR-00228748 at -48 (Within Incognito, "[u]sers who perform a Google search, it is logged against a Zweiback (lifetime [redacted]) that is identifiable by IP address"); GOOG-BRWN-00467569 ("Incognito mode … does not anonymize IP address despite a common misconception that it does"); GOOG-BRWN-00613802 ("[W]e actually keep all of these zwieback-keyed logs tied to IP address for [redacted]").

228. There are two versions of IP addresses in common use throughout the class period – IPv4 and IPv6.[103] IPv4 has 32 bits, which defines approximately 4 billion IP addresses (2^32). IPv6 has 128 bits, which defines approximately 3.4 x 10^{38} IP addresses (2^128) – equivalent to approximately 43 quadrillion-trillion IP addresses for every individual in this world. Google collects statistics about IPv6 usage and reports that within the United States, IPv6 adoption rate is approximately [redacted].[104]

229. Within Google's various logs and storage, IP address, particularly IPv6, can be used to link data from different logs storing signed-out data collected from private browsing and non-private browsing modes to a particular user's device and to the user's signed-in identity and data. As Google explains, "Access to third-party [non-Google] content (such as advertising) hosted on a page can be correlated by IP address. For instance, if ads hosted by adtech.com [for example, Google] are displayed on several sites (say foo.com and bar.com), adtech.com will be able to correlate those ad views across pages by source IP addresses. Furthermore, if the user visits adtech.com with the same IP address, any identity information from that visit (login, cookies) can be correlated with ad views on third-party [non-Google] sites, even if the browser is configured to block or periodically flush third-party [non-Google] cookies." (GOOG-CABR-04215325 at-328). During his deposition, Google engineer Michael Kleber described IP address as "one [] tool that could be used to recognize the same browser as it visits different websites" and predicted that as "third-party [non-Google] cookies go away," actors who like "collecting [] cross-site activity" might use IP addresses to do so (Kleber January 14, 2022 Tr. 74:17-75:8). And Brian Rakowski,

[103] https://www.google.com/intl/en/ipv6/statistics.html (Last accessed on April 11, 2022).

[104] *Id.*

the so-called "Father" of Incognito, admitted that IP address can be used for fingerprinting (Rakowski Tr. 50:16-18).

230. Different devices from the same household may share the same public IPv4 address. These devices can be differentiated by their User Agent – which identifies the type of device (Windows, Mac, Android, iPhone, etc.), browser (Chrome, Edge, Firefox, Safari, etc.) and version.[105] In addition, browsers such as Chrome use User Agent Client Hint (UACH), which specifies the browser brand name (e.g., "Google Chrome", "Microsoft Edge", etc.).[106] Under typical usage scenarios, it is practically impossible for two different users to have the exact same combinations of IP address and User Agent at all times during the class period. Google engineer Michael Kleber testified that "the user-agent string is automatically included in every request that the browser makes to a web server. And if your hope is limiting the amount of, like, finger-printable information that might be learned about a web browser, then that's sort of a ***worst-case scenario, right***? Here's a bunch of information, it might be different on [the] different browsers [of different users], and it's, like, automatically sent and given out to every web server that the browser talks to" (Kleber January 14, 2022 Tr. 161:3-13 (emphasis added)).

231. One metric for user identifiability is called entropy, measured in number of bits of data needed to uniquely identify a person. With a little less than 8 billion people on earth, 33 bits of

[105] GOOG-CABR-04639899 "User agents identify themselves to servers as part of each HTTP request via the `User-Agent` header. This header's value has grown in both length and complexity over the years… There's a lot of entropy wrapped up in the UA string that is sent to servers by default, for all first- and third-party requests. This makes it an important part of fingerprinting schemes of all sorts, as servers can [passively](https://w3c.github.io/fingerprinting-guidance/#dfn-passive-fingerprinting) capture this information without the user agent's, or more importantly the user's, awareness or ability to intervene or prevent such collection."

[106] https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Sec-CH-UA (Last accessed on April 11, 2022).

data is needed to uniquely identify a person (2^33 = 8.6 billion).[107] With around 330 million people in the United States, 29 bits of data is more than sufficient to identify a person (2^29 = 537 million). Google engineers measure that IP address alone contributes to [redacted] of entropy (GOOG-BRWN-00601937).

232. Because of its "highly identifying surface", Google engineers "propose[d] NOT recording the client's public IP address" (GOOG-CABR-00059774 at -780). Similarly, Google's Privacy and Data Protection Office proposed "masking IP addresses" within Chrome Incognito "to deliver stronger guarantees to users" such that "the user's behavior would be private with respect to their ISP and origins, as well as to Google to the extent possible" (GOOG-BRWN-00650016 at -16). Testifying about this document, Google engineer Michael Kleber described how a "goal" of this proposal was to "keep[] your activity from being observed by somebody who can watch the Internet traffic coming or going to your browser (Kleber March 18, 2022 Tr. 43:20-44:4).

233. However, Google has continued to collect and store IP addresses in logs, including IP addresses that Google collected throughout the class period from private browsing. Together, IP address and User Agent "carries [redacted]" of entropy (GOOG-CABR-04635379 at -380), which is more than sufficient to uniquely identify individuals in the United States. And this does not even account for cookies, screen resolution, and the myriad other information Google collects from its users and stores on Google's servers. Google's internal documents recognize that "[a] typical user's browser currently sends enough entropy to allow for fingerprinting, regardless of what GWS IDs we send" (GOOG-CABR-00487012 at -022). In other words, even without any Google identifiers, Google's logs contain enough high-entropy data to re-identify a person whose activity has been tracked, including those users who were in a private browsing mode, such as Chrome Incognito.

[107] GOOG-CABR-00800511 at -512: "[redacted] of information is enough to specifically identify a person"

234. As Google explains in GOOG-CABR-04417613 at -616, IP address is the join key across "Authenticated (i.e gaia)" and "unauthenticated (zwieback, biscotti…)" ID spaces. This is corroborated by the data obtained through the Special Master process. For example, GOOG-BRWN-00819830 shows Plaintiff Jeremy Davis's Google Account ID: [redacted] name, email address, IP address: [redacted] and user agent.[108] This same IP address and user agent are logged in his GAIA, and Biscotti (display ads) logs.[109] On a different device, Mr. Davis's Zwieback (search ads) and Biscotti logs containing Incognito traffic on non-Google websites while he is not signed-into his Google account also contain the same IP address and user agent.[110] As another example, the same IP address and user agent are logged in GAIA, Biscotti and Zwieback

[108] Several different user agent strings are associated with Mr. Davis' IP address [redacted] depending on the device he used and the type and version of the browser at the time of his logged activity. Each IP address and user-agent string pair is associated with a timestamp in GOOG-BRWN-00819830. For example, Timestamp: [redacted] IP Address: [redacted] Activity: [redacted] and User Agents: [redacted]

[109] For example, a preserved GAIA log: [redacted] contains a [redacted] field in row 2 equal to [redacted] This integer is converted to IP address [redacted] at https://www.browserling.com/tools/dec-to-ip. The same row contains "user agent: [redacted]" Produced from the First Iterative Search in the Special Master process, GOOG-BRWN-00826550 contains Mr. Davis' signed-out data in [redacted] associated with Biscotti ID [redacted] which contains the same IP address and user agent. For example, row 2 of GOOG-BRWN-00826550 contains IP address in RemoteHost: [redacted] and UserAgent: [redacted] The user agent corresponds to a Windows device and Chrome browser.

[110] For example, from the Second Iterative Search, production 2022-03-14 Brown v. Google - [redacted] – [redacted], and [redacted] all contain the same IP address in the RemoteHost field as [redacted] and the same UserAgent: [redacted] These logs contain data from Incognito sessions while Mr. Davis was not signed into his Google account. The user agent corresponds to a Mac device and Chrome browser. Additional support information can be found in Appendix H.8.

logs while Dr. Dai is in Incognito mode and not signed into Google.[111] Thus, while Google's Zwieback and Biscotti cookies are "pseudonymous" identifiers (*e.g.*, Kleber January 14, 2022 Tr. 47:18-48:15), it is possible to re-identify the user these IDs are associated with using IP address and user agent string, as demonstrated by the analysis I conducted in this case.

235. IP address is also the join key between private browsing and non-private browsing data. For example, Google engineers explain that "Currently we are logging all user activities in incognito mode server-side, and that is more or less linkable to users signed-in data" (GOOG-BRWN-00184875), "we already consider it possible for Google to join regular and Incognito sessions" (GOOG-CABR-00489377 at -384), and "IP can be used to join the authenticated and incognito sessions" (GOOG-BRWN-00157001); *see also* McClelland Tr. 212:13-212:24 ("Q. Do you still agree with the statement, it's possible for Google to join regular and Incognito sessions? . . . A. As far as I know, assuming nothing has changed, then, yes, it should still be possible.").

[111] For example, GAIA log from the First Iterative search [redacted] contains a [redacted] field in row 2 equal to [redacted] which is converted to IP address [redacted]. The same row contains user_agent: [redacted] Also from the First Iterative Search process, GOOG-BRWN-00826529 ([redacted]), GOOG-BRWN-00826530 ([redacted]), GOOG-BRWN-00826531 ([redacted]), GOOG-BRWN-00826532 [redacted]), GOOG-BRWN-00826534 [redacted]), GOOG-BRWN-00826535 ([redacted]), GOOG-BRWN-00826536 ([redacted]), GOOG-BRWN-00826537 [redacted] and GOOG-BRWN-00840745 [redacted] contain Dr. Dai's Incognito signed-out data associated with Biscotti ID [redacted] GOOG-BRWN-00826130 [redacted] contains Dr. Dai's Incognito signed-out data associated with Zwieback ID [redacted]. All of these Incognito signed-out Display and Search ad logs contain the same IP address and user agent as that in the GAIA log: RemoteHost: [redacted] or client_ips: [redacted] and UserAgent: [redacted] The user agent corresponds to a [redacted] Additional support information can be found in Appendix H.9.

236. Continuing with the previous example using Mr. Davis's data in Google logs, it is apparent that the same IP address and user agent appear in both Chrome Incognito logs and regular mode logs – permitting joinability and identification.[112]

237. I was not provided access to data concerning all class members, and the Special Master process only provided limited data access. Based on that analysis, it is clear to me that Google can link private browsing data and identify class members based on the data Google stores. One issue in that respect, however, is preservation. Had Google preserved IP addresses and User Agent strings at the beginning of this litigation, it would have been possible to use IP addresses and User Agent strings to join a user's private browsing activities on non-Google websites with the user's Google account identity for a longer period of time. However, I understand that Google did not preserve this information and that Google has continued deleting data even after this lawsuit was filed in June 2020.

238. Not only did Google collect sensitive and fingerprinting private browsing data throughout the class period, which can be used to join user data across ID spaces and across private browsing mode and non-private mode, it has also stored the data in a way that can be linked to a user's identity even if the user is not signed into Google. This is clear from Plaintiffs' data obtained through the Special Master process and Google's internal documentation. For example, in mid-2016, Google began a multiyear overhaul of their Display Ads system to enable the association of

[112] For example, from the Second Iterative Search, production 2022-03-14 Brown v. Google - [redacted] – AEO, [redacted] (search log) and [redacted] (display log) contain Mr. Davis's browsing data in Chrome regular mode while he was not signed into his Google account. Both logs contain the same IP address within the [redacted] field: [redacted]". This is also the same IP address shown earlier in his Incognito data logs from the Second Iterative Search. Furthermore, the same User Agent [redacted] appears in these Incognito and Regular mode data logs as well. Additional support information can be found in Appendix H.10.

user data to their Google account information. This effort was codenamed [redacted] (abbreviated as [redacted]).

239. Under the [redacted] architecture, Google chose to store encrypted Biscotti IDs within GAIA logs, with the decryption key being available for [redacted]; similarly, Google chose to store encrypted GAIA IDs within Biscotti logs (GOOG-CABR-04773853 at -888). As one example, if an Incognito user browses non-Google websites while signed out of Google, but later within the session signs into a Google account, then a subsequent ad request could "carry both the Gaia cookie and th[e] signed-out [Biscotti] cookie" (March 18, 2022 Berntson Tr. 113:2-9).[113]

240. Google engineers point out that this Google storage structure deterministically links GAIA and Biscotti (thereby linking a user's pseudonymous identity to their Google account identity): "narnia will happen at some point (linking gaia and biscotti)" (GOOG-BRWN-00161432 at -434) and "deterministic join between GAIA and Biscotti ID: After the [redacted] look back window a linkage between G->B can be obtained based on data that is < [redacted]" (GOOG-CABR-04724084 at -106). Google engineers also indicate that Google executives could authorize the Biscotti IDs stored in GAIA logs to be decrypted: "display ads personal logs contain encrypted biscotti, so you could technically see if the same biscotti exists in [both] personal and non personal logs" (GOOG-BRWN-00471401 at -401). Just because an ID is encrypted does not mean it is secure. Google acknowledges that "we don't yet have a crypto-erasure guaranteed key management system fully integrated with all ID joinability-sensitive encryption keys. This will be critical in our ability to guarantee the satisfaction of privacy safety constraints in terms of avoiding

[113] When asked during his deposition "would it bother me if I had learned that there was a pseudonymous cookie that was attached to the user's account," Martin Shelton, former Google user experience researcher, responded "I personally... wouldn't be in love with that." Shelton Tr. 100:9-16. With at Google, Mr. Shelton authored a document titled "Five Ways People Misunderstand Incognito and Private Browsing." *See* GOOG-BRWN-00061607.

risks due potentially to internal misconfiguration and/or external inquiries." (GOOG-CABR-04706890 at -893 (emphasis omitted)).

241. Through a series of slides in GOOG-CABR-04773853, Google engineers showed that Google's [redacted] architecture with GAIA and Biscotti IDs stored in the same log (called Option 2) had significant subpoena and joinability risks and an alternative implementation (Option 1) was markedly superior in these respects. However, in choosing between the more privacy preserving Option 1 vs. the less costly Options 2, Google chose to implement Option 2 irrespective of its privacy risks (GOOG-CABR-04724084 at -106).

242. Google's collection and logging practices create a risk for users that Google will be required to produce their private browsing activity in response to a subpoena. Google retains encryption keys for Biscotti identifiers from signed-out browsing in Google's logs for signed-in browsing ("p-logs") for [redacted] (GOOG-CABR-04773853, -854, -888). With those encryption keys, Google can decrypt Biscotti identifiers in personal logs, and then search the signed-out logs for that same identifier (GOOG-CABR-04773853, -854, -867). Because the encryption key facilitates this join between signed-out and signed-in data, Google would be required to produce signed-out data for a Google account-holder in response to a subpoena for that account (GOOG-CABR-04773853, -54; GOOG-CABR-03652549, -552). Google recognized the risk that its data collection and logging practices would require it to produce this signed out browsing activity—including private browsing activity (GOOG-CABR-03652549, -552-53). I have not found in any of Google's records from Google's production that indicate Google implemented double-logging for ads or another privacy feature that would prevent this outcome.

243. Google's [redacted] storage architecture has been demonstrated through Plaintiffs' data. Initially, Google asked the Special Master to withhold the production of a particular GAIA log

[redacted]). When data from this log was finally produced to Plaintiffs on February 23, 2022, it was apparent that data from this GAIA log can be associated with Biscotti IDs. In fact, Google named each produced file not by the GAIA ID but by the corresponding Biscotti ID.[114] Through a user's GAIA ID, Google can obtain the user's Biscotti IDs that were generated during the same browsing sessions, even if the user is signed out of Google during a portion of the same sessions.

244. This is another example of where the lack of preservation impacts the potential to identify class members individually. Although this can be done for some period of time, had Google preserved Biscotti encryption keys, which are used to decrypt the encrypted id,[115] it would have been possible to locate additional user's signed-out private browsing activities via their GAIA IDs for users who may have, on occasion, signed into their Google account in a private browsing session. However, Google did not preserve this information.

245. In addition to IP address, User Agent string and GAIA IDs, another way Google's logs link user identity to signed-out private browsing data is through signed-in identities on non-Google

[114] For example, from the First Iterative Search, production 20220223 Brown v. Google - personal-[redacted], GAIA log [redacted] contains "gaia_id: [redacted]". This GAIA ID belongs to Plaintiff Jeremy Davis as discussed in the running example. His Biscotti ID appears as [redacted] within the same log. The file name of the [redacted] GAIA log contains the decrypted Biscotti ID: [redacted]. GOOG-BRWN-00826401 contains Mr. Davis's signed-out data in [redacted] associated with the same Biscotti ID and the same IP address and user agent. As another example, GAIA log [redacted] contains "gaia_id: [redacted]" with the encrypted version of Biscotti ID [redacted] stored within the same log. This GAIA ID belongs to Dr. Dai. GOOG-BRWN-00826529 contains her Incognito browsing data in [redacted] associated with the Biscotti ID [redacted]. Additional support information can be found in Appendix H.9.

[115] GOOG-BRWN-00433264 at -272: [redacted] and GOOG-CABR-04773853 at -867 "Google could be compelled to decrypt all Biscotti IDs in P-logs records for this GAIA ID and then have to return all records (for all time) with those Biscotti IDs."

websites such as Publisher Provided ID ("PPID") and Analytics User ID.[116] Throughout the class period, when a user logs-in on non-Google websites that enable signed-in identifiers, whether in private browsing or non-private browsing mode and on any browser and device, the same identifier is associated with the data Google collects from a user's browsing activities on that website. Google further logs all such data within the same logs and uses these data for serving personalized ads. As a senior Google Ads engineer testified, "if the same PPID is supplied by the same publisher in a subsequent user interaction, the same mapping entry will be used to ensure the same PPID is mapped to the same Biscottis ID. . . . [O]nce the mapping entry has been created in subsequent times, [redacted] will be looking up that entry and therefore obtaining the same mapped Biscottis value for subsequent uses" (Liao Depo. 44: 22; 48:24 - 49:2). Google then stored these data in logs along with other detailed fingerprinting information tied to the users and their devices, including IP addresses, User Agent, cookies and time stamps.[117] Similarly, Google stores Analytics User ID and other identifiers such as Analytics CID and Biscotti cookies within the same logs.[118] Thus, through a user's signed-in identifiers on non-Google websites, Google can obtain the user's Biscotti IDs that took place in the same browsing sessions.

246. Using Mr. Davis' Analytics data from the Second Iterative search as an example, I have been able use his Analytics User ID to locate his Biscotti ID [redacted] generated

[116] Both PPID and Analytics User ID have been in use since at least 2014 (GOOG-CABR-03622146, GOOG-CABR-04691939 at -940).

[117] *See* Halavati Depo. 147: 22-25 ("Besides the ID and zwieback cookie we also record the browser site, which is the user agent" (referencing GOOG-CABR-00832014)); Liao Depo. 69: 18-20 ("I believe there generally is a timestamp associated with entries in the [PPID Biscotti] mapping table.").

[118] "2022-03-17 Brown - Letter to Mao re Chung Deposition Follow-up" acknowledge that the Analytics User ID, CID and Biscotti ID are stored together in two logs: [redacted] and [redacted]; *see also* Chung Depo. 113: 3-9 (stating that signed-in User ID can be stored in Analytics logs with client device ID).

from Chrome Incognito browsing. From the Biscotti ID, one can examine Google logs containing X-Client-Data or the maybe_chrome_incognito field to determine Mr. Davis' Chrome Incognito browsing activities.[119] In addition, I have been able to use the same Analytics User ID to locate Mr. Davis' Biscotti ID [redacted] generated from non-Incognito browsing; thus demonstrating that the same Analytics User ID can join data from Incognito and non-Incognito browsing.[120]

247. From the First Iterative Search in the Special Master process, I have observed PPID-mapped-biscotti associated with regular Biscotti (IDE) data. For example, GOOG-BRWN-

[119] From the Second Iterative Search, production "2022-03-25 Brown v. Google – Analytics [redacted] data – AEO", Google produced Plaintiffs' Analytics data. For example, file "By [redacted]" row 2246 corresponds to Plaintiff Mr. Davis' Analytics User ID [redacted] from [redacted]. Column M of the same row contains a request URL containing his Analytics CID from [redacted] This same CID is found in the file [redacted] for example, row 5 and column M. The same row, in column A contains Mr. Davis' Biscotti cookie [redacted]. The embedded Biscotti ID in this cookie is shown in "2022-03-02 Brown v. Google - Decode IDE.pdf", page 4, item 33 as [redacted]. I have discussed earlier that this Biscotti ID is associated with [redacted] containing data from Incognito sessions while Mr. Davis was not signed into his Google account. The same Biscotti ID is associated with other ads logs such as [redacted] As of April 11, 2022, Google has yet to produce Plaintiffs' full data from the [redacted] logs containing the maybe_chrome_incognito field and other ads logs requested for the Second Iterative Search.

[120] For example, file [redacted] row 2237 corresponds to Plaintiff Mr. Davis' Analytics User ID from [redacted]. Column M of the same row contains a request URL containing his Analytics CID from [redacted]. This same CID is found in the file [redacted] for example, row 131 and column M. The same row, in column A contains Mr. Davis' Biscotti cookie [redacted]. The embedded Biscotti ID in this cookie is shown in "2022-03-02 Brown v. Google - Decode IDE.pdf", page 5, item 38 as [redacted]. I have discussed earlier that this Biscotti ID is associated with [redacted] containing data from a Chrome non-Incognito session.

00840745[121] contains Dr. Dai's Incognito data associated with Biscotti ID [redacted] in the [redacted] log. Row 208 in GOOG-BRWN-00840745, corresponding to a single entry in the log, contains 20 page of data (See **Exhibit C**), among which is a [redacted]. This indicates that PPID (which identifies a signed-in user on non-Google websites) is joinable with the same user's private browsing information associated with regular Biscotti IDs in the same way Analytics User IDs are. As of April 11, 2022, Google has not produced data from ads logs containing PPID related information from the Second Iterative Search and subsequent searches. Google has also not produced full data from logs containing "maybe_chrome_incognito", nor any data from [redacted] logs containing "is_chrome_incognito" and "is_chrome_non_incognito" fields. I reserve my rights to supplement this report should Google produce the requested data later.

248. Had Google preserved users' signed-in IDs on non-Google websites, which uniquely identify a user, as well as associated logs, it would have been possible to locate additional records of user's signed-out private browsing activities. However, Google did not preserve this information.

249. In addition to storing IDs from different ID spaces in the same log, which allows data to be joined, Google also includes IDs from different ID spaces, such as "authenticated" IDs, "unauthenticated" IDs and signed-in IDs on non-Google websites, within the same URL link. Through the Special Master process, I have observed that embedded IDs within URL links, which Google receives along with intercepted private browsing data, have been decrypted.[122] This reveals

[121] GOOG-BRWN-00840745 was produced as a part of the First Interactive Search on Jan. 31, 2022 in PROD71 containing data in [redacted] logs.

[122] 2022-03-14 Brown v. Google - Dycrypted URLs - AEO.xlsx

a single URL link carrying both GAIA and encrypted signed-out ID. In addition, PPID, Biscotti ID and an Ad Manager first party cookie called GFP cookie are also carried in the same URL. Other ID combinations are also possible.[123]

250. These IDs, along with other cookies and information intercepted and transmitted to Google, reach Google servers and are accessed by a single entity called the [redacted] Google describes [redacted] (GOOG-CABR-00543864 at -872). [redacted] (GOOG-CABR-00543864 at -876 and -877) [redacted] (GOOG-CABR-00543864 at -884). [redacted] (GOOG-CABR-05876612 at -612).

251. Yet another example of joinability is through common data and information stored in the various logs, such as common identifiers and event IDs.[124] For example, search and display logs are joinable through common identifiers stored within, such as the advertiser user ID (AUID). As Google's internal document explains, "AUID causing cross-id-space joinability: Each AUID can

[123] GOOG-CABR-04126517 at -517: "User IDs are an important part of an ad serving system. Despite the simplified view that each request has a single user identifier, we have a great deal of complexity when it comes to determining user identifier(s) for each request. Each request may come with one or more IDs encoded in a cookie or sent as a request param. We call these request IDs. Examples include biscotti id (from biscotti cookie), DFP PPID, iOS IDFA, and Android Advertiser ID. In some cases, we map a request ID to a canonical Google ID at server side. Examples are mapping a mobile device ID (such as IDFA) or PPID to a server-side biscotti ID."

[124] For example, GOOG-CABR-03653330 at -351 and -353 describes Event IDs and other information stored in logs as "Potential deterministic join key".

be mapped to multiple Gclids. And the Gclids can be associated with different id types. Gclid is only encrypted for signed-in clicks, so even without decryption this could lead to Biscotti-Zwieback joinability, if the user has both Display and Search clicks" (GOOG-CABR-04076570 at -573). Indeed, Dr. Glenn Berntson testified that an AUID is mapped to a click ID, which itself can be mapped to a Biscotti or Zwieback (March 18, 2022 Berntson Tr. 83:19-84:4).

252. For private browsing data that Google has already deleted from its logs, had Google preserved a subset of that data, including the various identifiers, their encryption keys and linkage/mapping tables, it would have been possible to link users' signed-out private browsing activities with their signed-in identity on non-Google websites or with their search activities.

253. In addition to storing private browsing data in a way that can be joined with a user's identity -across search and display logs, Google also proactively connects collected data under a "User ID Relation Graph" for conversion attributions and other use cases. Google explains that "User Id relation graph is a graph … that relates user-id with other identifiers (device-id physical, aaid, idfa, phone, email)" (GOOG-CABR-04782308 at -309). As the following figure shows (GOOG-CABR-04732430 at -434), the User ID Relation graph links personally identifying information with signed-out IDs, which admittedly presents a "privacy problem."

254. Google also maintains a "User Trust Graph" in connection with Publisher and Advertiser data as shown in the Figure below (GOOG-CABR-04732430 at -445). "The user id's are nodes of the graph. Aggregated information about user activity and information about the user is anchored on the node as an attribute" (GOOG-CABR-04782100 at -101). This graph effectively combines signed-in and signed-out data. Google considers this graph to not increase its privacy risk since the "AdSpam team already has event level logs access across id spaces" (GOOG-CABR-04732430 at -433). Thus, interconnecting user data is simply a matter of authorization control within Google. It is not a question of whether data is joinable (it is), but simply a matter of who within Google can have access to the joined data.

255. Another significant source of fingerprinting data Google collects is User Metrics Analysis ("UMA") data. UMA is a built-in feature in Chrome that reports extremely detailed Chrome usage information to Google, including Incognito mode usage. This includes highly sensitive and fingerprinting information[125] that is sent to Google every [redacted] from desktop devices and even more frequently on mobile devices[126]:

- [redacted]
- [redacted]
- [redacted]
- [redacted]
- [redacted]

[125] GOOG-BRWN-00032906, GOOG-CABR-00000015, "2021-11-18 Brown Omnibus Sheet.xlsx", and Plaintiff's data.

[126] GOOG-CABR-04970427 at -433: "UMA information is sent every [redacted]" and GOOG-CABR-04899841 "for mobile it is [redacted] based on if user is on wifi or cellular."

[127] GOOG-CABR-00374449 showing UMA demographics data launched in 2019 and GOOG-CABR-03710192 at -193: "increasing the appeal of Chrome to Gen-Z users has been identified as an important goal for Chrome. The demographics data in UMA will be critical in that effort because it will allow us to conduct experiments targeted at Gen-Z users as well as track their usage of various features of Chrome."

256. Prior to mid-2016, users had to opt-in to report UMA information; however, Google changed this to opt-out in 2016 to gain "competitive edge" over its competitors.[128] After this change, Google saw UMA participation rates increase significantly.[129]

257. UMA data is highly identifying and can be joined with other data logs containing user's private browsing information. Google engineers recognize that storing so much identifying information with a stable UMA ID is a joinability risk and urges restraint in data collection.[130] As

[128] GOOG-CABR-03710213 at -213: "Currently, across all platforms, users have to explicitly opt-in to send usage metrics to Chrome. These usage metrics are critical for teams to understand both how users are using Chrome and how Chrome is performing in the wild… By moving UMA to opt-out, we hope to reduce our bias by capturing metrics for a much greater set of users, particularly those who are willing to share data but don't explicitly opt-in in the current system."

[129] See e.g, GOOG-CABR-03717199 at -201: UMA opt-in rate at [redacted] on Windows and Mac and [redacted] on Android and iOS in 2014. And GOOG-CABR-04743125 at -126: % of users not opt-out of UMA being [redacted] on different platforms and [redacted] on Android in 2021.

[130] GOOG-CABR-04293141 at -143: "I'm wondering, how do you think about the problem of stable (hardware) entropy, specifically on Chrome OS where we collect rich hardware information (e.g. [redacted], ...)? Should we limit the amount of hardware configuration information we collect? … Personally, I'm not happy with the situation we're in, and I wonder whether we could evolve UMA in a direction that reduces joinability risks. Concretely, I believe a large part of the problem stems from the large amount of data keyed to a single, stable client id. … I do understand that there are significant advantages to a monolithic data set from the data analysis point of view, but this same property that enables a wealth of analysis, also seems to

one Google engineer explained, "the main concern is making it difficult to join UMA and Google logs. I agree … that fundamentally this is a policy issue. Indeed, I think joined these data sets can already be done by malicious insiders with access to the underlying logs ... I think it's likely that by looking at a fingerprint of activity--what sites are visited which days--for a particular client, it will be possible to join that client ID to Google logs. In short, I agree the statement that developers should not join UMA and Google logs is a policy. From the technical side, it's likely already possible" (GOOG-BRWN-00232201 at -202 and -203).

258. Another Google document acknowledges that it is possible to join a UMA ID with a user's Google account GAIA ID using omnibox search user action and timestamp.[131] This is significant since a user's Google account GAIA ID uniquely identifies a class member across any of a user's devices, old or new. Even if a user can no longer locate their UMA ID in an old device that they no longer possess or if the user factory resets a device, it is possible for Google to find a user's UMA ID by correlating the user's Gaia logs and UMA logs with search actions and timestamps. Knowing that detailed event level data are stored with timestamps, IP address and user agent in these logs, it is my opinion that UMA can be joined with ads event logs.

preclude others because (in my opinion) we need to show restraint with regards to the amount of hardware information that we can include." *See also* GOOG-BRWN-00699534 ("we can consider a scenario w[h]ere each piece of UMA stats collected is not identifying user, but putting them all together gives a good sketch.")

[131] GOOG-CABR-00430662 at -662: "UMA/Gaia join: some worry that experimentids could allow a join between a UMA-users Finch (Chrome install) id and a Gaia id. In fact it is probably easier to do this join today with the timestamps of omnibox interactions which result in a search."

G. California: Google designed its systems to provide its California-based employees with access to private browsing information collected nationwide, and Google employees routinely access that information in California

259. Throughout the class period, Google employees in California have had ready access to the private browsing data at issue in this lawsuit. Google designed its servers in a way that they are in various locations but with access granted to its employees in its headquarters in California. Google has admitted that "Google employees, including those based in California, may access user browsing data provided they have proper access permissions" (RFA No. 44). And Google has agreed to amend its response to RFA No. 44 to clarify that this "user browsing data" includes data collected by Google from users' visits to Google Analytics and Google Ad Manager, which would include the private browsing data at issue in this lawsuit, collected across both classes and throughout the class period.

260. In my experience, Google's server strategy is typical. Many tech companies have servers located in various locations, but with access to the data from their headquarters. This is in part to protect against power outages or other events that can impact specific geographic areas, and in part to avoid network congestion by placing servers closer to the markets they serve. Google has similarly designed a way to store data in various locations but to ensure access in California. This has been the case for the entire class period, and across the various sets of data at issue in this lawsuit.

261. For example, Dr. Berntson's deposition testimony shows that Google stores data in ways that enables access to data in California. Dr. Berntson testified that "there is at least one data center in California" (Berntson Tr. June 16, 2021, 41: 14-20). Google engineer Chris Liao testified that Google's data storage are accessible by employees in California (Liao December 3, 2021 Tr. 14:8-17:8). Furthermore, both Chris Liao and Bert Leung testified that they are located in California and that their teams in California have access to at least servers related to display ads and ads

identity (e.g., [redacted] and the [redacted] which process all display ads traffic, including traffic from private browsing).[132] And as explained by Dr. Berntson, when an ad request comes into Google's systems, "[redacted] analyzes all of this incoming data, and it's the thing that basically sets all of the rules in terms of whether downstream systems can access and use user identifiers" (Berntson March 18, 2022 Tr. 95:21-97:15).

H. Class Member Identification: Throughout the class period, Google collected private browsing information in ways that can be used to identify class members, though Google also withheld and destroyed data relevant to that identification

262. Through its tracking beacons, and throughout the class period, Google collected private browsing information and stored this information in numerous Google logs and data repositories, including IP address, device/browser type and version, cookies and IDs, date and time, and websites visited as I discussed in Section D of this report. In addition, Google collected detailed Incognito usage information as well as used Incognito signals in live traffic to determine and log Incognito usage. Based on my analysis of the discovery in this case, it is my opinion that such information can readily be used to identify individual class members. However, it is my understanding and opinion that Google also withheld and destroyed relevant data that could be used as part of that identification, including during the discovery process. The data destroyed by Google could have been used to identify additional class members.

[132] Leung Tr. 82:21-25 (works in California), *id.* 18:24-25 ("UIS is part of my responsibility"), *id.* 263:14-16 ("I am part of the ads infrastructure team, which part of our responsibility is . . . ownership of [redacted], develop and maintain it"); *id.* 41:12-19 ("I am in the team who is responsible for building a dashboard to monitor certain browser for some of Chrome . . . ."); *id.* 16:19-21 (supervised by Chris Liao); Liao December 3 Tr. 15:18-16:1 (works in California); Liu Tr. 16:13-14 (Bert Leung and Chris Liao knew about the maybe_chrome_incognito field), *id.* 21:23-25 ("I was working with Bert and Chris on this third-party cookie blocking browser project"), *id.* 23:9-10 ("I was told by Bert to do this project"), *id.* 46:22-47:1 (Bert Leung is in charge of approvals for using the maybe_chrome_incognito field); Factual Stipulation Re: maybe_chrome_incognito (confirming that Bert Leung made the decision to implement the maybe_chrome_incognito field within [redacted] Google logs, and that the third-party cookie blocking dashboard uses this field); Porter-Felt November 16, 2021 Tr. 9:17-19 (business address is California)

263. I understand that the Court has scheduled an evidentiary hearing to address certain allegations regarding Google's discovery conduct and that the Court may make certain findings of fact and conclusions of law in connection with that evidentiary hearing. I provide the below opinions and information based on my current understanding. As noted previously, I reserve the right to update my opinions based on any subsequent developments, including with respect to that evidentiary hearing and any future findings or rulings from the Court.

H.1 Google's Lack of Production and Preservation

264. Google's production is deficient in several aspects as I describe in this section.

H.1.1 Lack of Production

265. Google provided a list of over 200 Google employees with relevant information to this case (GOOG-BRWN-00023909), but Google left out three employees whose work on Incognito detection are important to this case. Through persistent discovery efforts (e.g., requesting hyperlinked documents, briefing the need for additional document production, and deposition testimony), it has gradually and eventually come to light that Google engineers Quentin Fiard, Bert Leung and Mandy Liu are key Google employees who developed and implemented Incognito detection capabilities on Google's server side.[133]

266. On March 4, 2022, during Mr. Leung's deposition, Plaintiffs learned that his Incognito detection logic was implemented in Google's frontend server called [redacted] in a module called [redacted] and that logic was used by Google on live traffic (including based on the browsing activities of Class Members) to generate Incognito detection

[133] Leung Depo. 47:13-17 ("My role in this doc is to come up with the heuristic approximation of -- I mean, part of that is to come up with the heuristic approximation to -- for identifying whether the request may be coming from Chrome incognito"); *id.* at 76:19-77:6 (identifying Mandy Liu as the engineer who implemented the browser detection dashboard and Bert Leung as Mandy Liu's title manager). Sadowski Tr. 87:20-23 (identifying Quentin Fiard as the engineer "who knows about [redacted] logs").

Dashboard metrics (*See* Leung Depo. 186:25-187:6, 195:4-6).[134] On February 18, 2022, Google produced a document containing a vague reference by Mr. Leung to Ms. Liu in January 2021 stating that "You can reference how we did that before in [redacted]" (GOOG-BRWN-00845569). It now appears that Google has been actively using live traffic for Incognito detection for some time now since Mr. Leung's initial Incognito detection study in the spring of 2020. This was followed by Ms. Liu's implementation for Google to log the "maybe_chrome_incognito" field in various Google logs, including for display and search ads. Yet Google did not disclose these key witnesses to Plaintiffs.

267. Google has failed to identify many relevant log sources in connection with the Special Masters process. Google has not responded to Plaintiffs' questions as to what other logs may contain fields with "incognito" in fieldnames. Similarly, Google has not responded to Plaintiffs' questions as to what logs contain X-Client-Data header information. Even for some of the logs that Google did disclose, Google did not provide fields contained in the logs nor field descriptions for most of the logs. See Appendix E for a list of logs Google has disclosed to date and additional logs that Google did not disclose that may contain private browsing information.

[134] *See also* Leung Depo. 183:25-184:15 (testifying that "in the previous [redacted] analysis, there were an online monitoring dashboard that was trying to use the information passively received in the [redacted] server to compute that heuristic approximation maybe_Chrome_incognito bit. . . . [H]ere the [redacted] is running both, which as I said before, is one of the Google ads front-end server which means it -- when they serve -- where they are receiving certain requests from browsers, mainly Chrome, they may have access to their user agent or X-Client-Data header"); *id.* at 184:16-185:1, 185:7-8 (describing text reading, ""First check, from a Chrome browser; 2, Exclude Android WebView; 3, Exclude iOS"; and "4, Check X-Client-Data"" header as "path logic being add to [redacted] previously to -- for a computation of that approximate heuristic signal" and that, "as far as [he] can remember, those logic was added during the [redacted] analysis.")

H.1.2 Lack of Preservation

268. Google initially identified [redacted] log sources as relevant to this case – some with retention period of days and others, months.[135] On March 10, 2022, Google disclosed [redacted] additional logs, three containing an "is_chrome_non_incognito" field and [redacted] containing an "is_chrome_incognito" field. These Google logs have retention periods ranging from [redacted] to [redacted]. Three days before the close of discovery, Google further identified [redacted] logs and identified these, as well as two previously identified logs, as containing the "maybe_chrome_incongnito" field. My understanding is that Google has not changed the retention period of any of these log sources for the purpose of this litigation. Instead, in March 2021, long before Plaintiffs learned about these logs and their Incognito detection fields, Google took the position that it would be unduly burdensome for Google to "suspend its regular retention policies on the data stored in logs that record information received when a user visits a website that employs Google Ad Manager or Google Analytics services" (Dkt. 119 at 3).[136] I understand that Google did not at that time disclose to Plaintiffs that Google has already been logging the "is_chrome_incognito" and "is_chrome_non_incognito" fields since 2017 (Sadowski Tr. 70:11-18) and was in the process of developing and implementing the "maybe_chrome_incongnito" field in [redacted] logs.

269. Google's cookies and encryption keys for ID linkage purposes also expire after a preset number of days or months.[137] Also, IP addresses may be redacted from certain logs after several

[135] Data Source Information Provided Reconciliation - Brown.xlsx

[136] *See id.* ("Google estimates that suspending preservation of these logs—even if possible to do safely, without jeopardizing the data or Google's systems—would require storing over a [redacted]. Even if it were feasible, suspending the regular retention periods for the Disputed Logs and safely hosting the ever increasing data would take [redacted]").

[137] For example, Biscotti cookies expire in [redacted] or [redacted] (GOOG-CABR-04206179 at -182) and the encryption key for the Biscotti ID in GAIA logs expire in [redacted] (GOOG-CABR-04773853 at -888).

months (GOOG-BRWN-00029002). While Google has enriched itself with the collected data prior to their expiration or redaction, it appears Google did not make any effort to preserve this information for the purpose of class member identification in this case.

270. On or about February 28, 2022, I first learned that a set of signed-in data was preserved for the named Plaintiffs. On March 2, 2022, Google admitted that signed-in data in [redacted] different signed-in GAIA logs were preserved.[138] There was no explanation as to why a corresponding set of signed-out data was not similarly preserved for the Brown Plaintiffs. As I have explained in Section F of this report, Biscotti IDs used in the same browsing sessions as the GAIA IDs are stored in GAIA logs. Google can decrypt the Biscotti IDs at the time of GAIA data preservation to similarly preserve data associated with the Biscotti IDs. Furthermore, GAIA logs contain IP address and user agent information, which may be used to identify signed-out logs containing the same IP address and user agent for preservation.

271. Google could have preserved a set of data for the purpose of identifying class members in this litigation, but Google did not.

- UMA data: Google represented that the raw UMA log containing UMA user-action and histogram data is retained for [redacted].[139] However, Google's internal document indicates that Google retained this log indefinitely prior to early 2017 and changed its retention policy to [redacted] only towards the end of 2017.[140] Certainly, this means that Google could have retained this log indefinitely for class identification in this lawsuit; however, they did not.

[138] 2022-03-02 List of Preserved Data Sources.pdf

[139] Data Source Information Provided Reconciliation - Brown.xlsx

[140] GOOG-CABR-04801490 "We're planning to change the retention of Chrome UMA weblogs (raw logs) to be [redacted]. This will specifically affect [redacted] log source & type. You might remember seeing a similar email earlier in the year - where we changed retention to the current value of [redacted]. ([redacted].)".

Further, Google implements a sampling approach for UMA data collection. Google explains that "[t]he metrics team controls how many of the consenting users actually upload data to Google. Currently (Dec 2019) these are roughly: Android - [redacted] Windows - [redacted] Mac - [redacted] iOS - [redacted] (GOOG-CABR-04477378 at -381). Not only did Google refuse to retain collected raw UMA information for the entire duration of this lawsuit (deleting records past 1.5 years), it also did not change its sampling policy to retain key UMA information for US Chrome users for the duration of this lawsuit. At a minimum, Google could have collected UMA client_id2 from all US Chrome users, IP address, device information and the few user actions and histograms it uses to compute Incognito statistics. For example, Google could have preserved just the user actions within Incognito window open and exit events; as well, Google could have preserved just the histogram actions related to page load counts in Incognito and non-Incognito modes. Google knows what these metrics are and their importance to class identification; yet Google did not collect or retain these data.

- [redacted] Logs: Google engineer Dr. Sadowski testified that the [redacted] search location logs have been logging "is_chrome_non_incognito" and "is_chrome_incognito" since 2017 (Sadowski Tr. 70:11-18). The default retention period of these logs are just [redacted].[141] These location logs contain IP address information and other identifiers (GOOG-BRWN-00848368) that could be used to identify class members along with the Incognito signals. Google could have preserved these logs for class identification, but Google did not.

- Incognito Detection Ads Logs:
    - As discussed earlier, Google has been transmitting an X-Client-Data header in Chrome regular mode and not in Incognito mode. Google logs X-Client-Data information in certain ads logs, including at least [redacted] and [redacted], both of which were used by Google engineers in their log-based Incognito detection study in 2020.[142] It is my understanding that Google has not provided information regarding the default retention period for [redacted]. The [redacted] log has a default retention period of [redacted].[143] Google could have preserved these logs, along with other logs containing X-Client-Data information for class identification, but Google did not.

[141] FINAL SADOWSKI REFERENCE SHEET

[142] E.g., GOOG-BRWN-00204684 at – 687, titled "How to Detect Chrome Incognito Mode?" and GOOG-CABR-05280756 "While UserAgent is generally logged, HttpHeaders are not (It's only logged in a few log sources). In go/[redacted]-monitor, I am using [redacted] and [redacted]."

[143] Data Source Information Provided Reconciliation - Brown.xlsx

In addition, both of the aforementioned logs contain an "is_chrome_non_incognito_mode" field.[144] This particular field is in what Google refers to as [redacted], which is a data structure that can be shared among numerous logs as listed in GOOG-CABR-04408590.[145] Thus, even logs that do not contain X-Client-Data can be populated with the "is_chrome_non_incognito_mode" field for class identification purposes, including [redacted], which stores display ad request information.[146]

- Since at least July 4, 2021, Google has been logging a "maybe_chrome_incognito" field into [redacted] temporary ads logs.[147] [redacted] of these logs have a default retention period of just [redacted]. Google has not informed the Plaintiffs the default retention period of all of rest of the [redacted] logs. Google could have preserved these logs since July 2021 for class identification, but Google did not.

- Logs containing user sign-in identifiers on non-Google websites:
  - Analytics User ID: Google has represented that the Analytics User ID, which identifies a signed-in user on particular non-Google websites, is stored in several Analytics logs.[148] Many of these logs also contain Biscotti IDs as well as IP address and User Agent.[149] The default retention period for these Analytics logs ranges between [redacted].[150] As I have shown in Section F, identifiers in these Analytics logs may be used, in conjunction with the Incognito Detection Ads Logs listed above to identify class members. Google could have preserved these logs for class identification, but Google did not.
  - PPID: As discussed earlier, PPID identifies a signed-in user on particular non-Google websites. Through the Special Master process, I learned of PPID-mapped-

[144] 2022-03-10 Brown v. Google – Fields for [redacted] Logs - AEO.xlsx

[145] The file name of GOOG-CABR-04408590 is "[redacted] reimport list for [redacted] and GOOG-CABR-00059481 at -531 explains that "[a] proto is a way we represent data at Google! It's essentially just a structured array, an element in position N should always mean the same thing."

[146] For example, GOOG-CABR-04737403 at -414 explains [redacted] has a [redacted] field to log [redacted] specific data. [redacted] log list. Most [redacted] flows log [redacted]. (CAT2|XFP|XFA) [redacted], etc."

[147] STIPULATION REGARDING GOOGLE'S "MAYBE_CHROME_INCOGNITO" FIELD

[148] 2022-03-17 Brown - Letter to Mao re Chung Deposition Follow-up.pdf

[149] See e.g., Plaintiffs' data in 2022-03-25 Brown v. Google – Analytics [redacted] data – AEO

[150] Data Source Information Provided Reconciliation - Brown.xlsx

biscotti ID in the [redacted] log along with other identifiers such as Biscotti ID.[151] This log has a default retention period of [redacted].[152] Google could have preserved this and other logs containing PPID or PPID-mapped-biscotti for class identification, but Google did not.

- Other logs containing user sign-in identifiers on non-Google websites: Google has not identified other logs containing other user sign-in identifiers. Dr. Berntson testified that there is at least a CRM-ID or Google Ads User ID that also identifies signed-in users on non-Google websites. *See* Berntson March 18, 2022 Tr. 118:6-11, 130:5-7, 131:10-132:1 Google could have preserved these logs for class identification, but Google did not.

- Data in Logs: Not all ads and Analytics data is needed for class identification. Only a small subset of fields is needed including (1) identifiers, 2) IP address and user agent, (3) HTTP header containing X-Client-Data, (4) URL, (5) date and time, and (6) "incognito"-titled field (e.g., "maybe_chrome_incognito") or field containing Incognito-detection (e.g., "browser_info_state" fields containing "maybe_chrome_incognito"). Google knows which of its logs contain Incognito signals (logs containing the X-Client-Data header or Incognito detection fields such as "maybe_chrome_incongnito", "is_chrome_incognito", and "is_chrome_non_incognito_mode") and could have preserved just those logs for the purpose of Chrome class identification in this case.

- IDs: Google knows which of its logs contain ID linkages and mappings or join keys and encryption keys useful for class identification purposes and could have preserved those data.

H.2 Google's Obstruction Throughout the Special Master Process

272. Google obstructed the data production being overseen by Special Master Brush. In my opinion, the most egregious obstructions include: 1. Not identifying certain logs; 2. Producing log schemas that omit key Incognito detection fields; 3. Not formatting search queries and search parameters needed to produce data; 4. Delaying Google's productions; and 5. Failing to preserve

[151] For example, data from the First Iterative Search in GOOG-BRWN-00840745.

[152] Data Source Information Provided Reconciliation - Brown.xlsx

Biscotti data tied to the Plaintiffs' GAIA data. Google's obstruction has made it more difficult for me to obtain relevant data.

273. First, Google did not identify to Plaintiffs and the Special Master numerous Google logs that Google's own employees used to analyze Incognito traffic, as well as logs that contain fields dedicated to detecting Incognito traffic. As I discuss in Appendix G, in 2020, Google engineer Bert Leung undertook a log-based analysis of Incognito traffic, aiming to identify and quantify Incognito traffic within Google's advertising logs. Mr. Leung specifically focused on [redacted] logs: [redacted] (Google's Response to Interrogatory No. 34) and March 17, 2022 Supplemental Response to Interrogatory No. 34). Yet Google omitted the latter four logs from its submission in response to the Court's September 17, 2021 Order for Google to "identify to the Special Master and Plaintiffs all databases and data logs (collectively, 'data sources') that may contain responsive information" (Dkt. 273).

274. Likewise, Google had been logging a "maybe_chrome_incognito" field in [redacted] ads logs since at least June and July 2021. [redacted] of these log sources had not been previously identified by Google. Google had also been logging "is_chrome_incognito" and "is_chrome_non_incognito" fields in [redacted] location logs since at least 2017. As Google's letter on April 1, 2022 shows, the "is_chrome_incognito" field is used to represent "if an entry comes from a Chrome web browser in the incognito mode" and the "is_chrome_non_incognito" field is "True if the event originates from Chrome in non-Incognito mode."[153] Google did not identify any of these [redacted] logs in response to the September 17, 2021 order.

[153] 2022-04-01 Brown v. Google - Google Letter.pdf

275. Nor did Google supplement its production to include any of these missing logs even after the Court issued another order on November 12, 2021, requiring Google to "provide a declaration, under penalty of perjury from Google, not counsel, that 1. To the best of its knowledge, Google has provided a complete list of data sources that contain information relevant to Plaintiffs' claims" (Dkt. 331, Ex. 1). This Order noted that one of the purposes of the Special Master process is provide "Plaintiffs the tools to identify class members using Google's data" (*Id.* at 4). I would have accordingly expected Google to identify data sources that its own employees had selected to use for the purpose of identifying and quantifying Incognito traffic, as well as the data sources where Google had chosen to implement an "Incognito" detection field.

276. As a result, with my involvement and input on the Special Master process, Plaintiffs did not select any of the logs containing "Incognito" in field names nor the [redacted] additional logs Google identified on March 17, 2022 that Google reviewed during their log-based Incognito detection study for the First Iterative Search of the Special Master process. Had Google timely identified all of the logs that Google studied as well as the logs with "Incognito" detection fields, and produced complete schema for them, I would have advised Counsel to select those logs for the First Iterative Search.

277. Second, for the logs that Google did identify, Google omitted the incognito detection fields from the log schema before producing them to Plaintiffs and the Special Master. For example, logs named [redacted] and [redacted] contain the "maybe_chrome_incognito" field, but the schema for these logs that Google produced to Plaintiffs and the Special Master omitted this field.[154] Had the schema been properly produced by Google with the "maybe_chrome_incognito" field, I would have advised Counsel to select these logs for

[154] See schema from the "Brown - 3.8-3.9 Field Names" production.

Search 1. But because Google did not provide an accurate and complete schema, Plaintiffs did not select these logs for the First Iterative Search. In fact, for the First Iterative Search, the effect of Google's obstruction was that Plaintiffs did not select (and therefore did not receive data from) any log that contains the "maybe_chrome_incognito" field. While the Plaintiffs selected one of the logs containing the "maybe_chrome_incognito" field for the Second Iterative Search, Google has yet to produce complete data for that log.

278. Similarly, when Google produced schema for [redacted] and, pursuant to a Special Master directive, produced a schema for [redacted] in December 2021, both schema omitted an "is_chrome_non_incognito_mode" field. I did not learn that these logs contained this field until on or after March 10, 2022, when Google made a supplemental production showing more fields in the schema.

279. Google's excuse for omitting these fields from the schema is that its [redacted] tool only includes any given log's top-100 fields for a schema. This excuse is implausible for a few reasons. First, as I discussed in Section D of this report, Google produced more than 100 fields for the [redacted] and [redacted], both of these logs are stored in a data repository called [redacted], just like the other ads logs for which Google only produced the top 100 fields.[155] Second, the fact that Google later produced supplemental schema for the [redacted] log and the [redacted] log shows that Google was at all times capable of producing schema with more than 100 fields.[156] It is my opinion that Google could have readily produced complete schema for all of these logs initially. At the very least, Google could have added the "Incognito" detection

[155] GOOG-BRWN-00029378 contains some of the logs that are stored in [redacted] In production "2022-03-25 Brown v. Google – [redacted] data – AEO", Google lists both [redacted] and [redacted].

[156] 2022-03-10 Brown v. Google – Fields for [redacted] Logs - AEO.xlsx

fields to their production, and had Google done so, Plaintiffs would have learned about the "Incognito" detection fields far sooner.

280. In short, Google undermined the Special Master process from the very beginning by withholding important logs altogether and redacting fields from schema spreadsheets, which made it far more difficult to select logs for searches. And as a result of Google's obstruction, I am still lacking data from the [redacted] logs that contain the "maybe_chrome_incognito" field or the "is_chrome_incognito" or "is_chrome_non_incognito" fields.[157] I likewise have not received data from the additional [redacted] logs that Bert Leung studied for his Incognito-detection analysis. Finally, it is my understanding that Google has yet to confirm whether there are any other field names within Google logs that contain "Incognito."

281. Third, Google undermined the Special Master process by not properly formatting search queries and search parameters. For the First Iterative Search, Google did not disclose that Analytics CID and UID searches require a corresponding web property ID. Neither did Google identify ways for extracting PPID. As a result, when search results were either empty or non-sensical, it was impossible for me to assess what was the cause. For the Second Iterative Search, Google introduced an error into the search that Plaintiffs proposed for IP addresses and user agents. Google modified the user-agent search term that Plaintiffs proposed by changing a semi-colon to an apparent hex value, which resulted in no results being returned.[158] Further, Google did not append a ",gzip(gfe)" string to the end of user-agent search term as these strings are stored in Google's

[157] Google has produced a limited set of data for some of the [redacted] logs containing "maybe_chrome_incognito". These include a limited set of live demo data (2022-03-04 Brown v. Google - Live [redacted] Queries.xlsx) and data in selected ten fields as a way for Plaintiffs to understand the content of some of these fields, given Google has not produced any field descriptions for these logs (2022-03-18 Brown v. Google - Field Values.xlsx).

[158] 2022-03-17 Brown v. Google - UserAgent_IP Search Script – AEO.pdf

logs.[159] I would expect Google to know how to search its own data sources and to craft searches to ensure they return responsive results. But Google failed to do so. I understand that Google will correct this error in later productions, but not in time before this report is due.

282. Fourth, Google undermined the process through its delay. According to the November 12 Order, Plaintiffs were entitled to conduct four separate, iterative searches (Dkt. 331, Ex. 1). Yet Google completed just one of those searches before it was time for me to submit this report. Under the order, Google should have completed the First Iterative Search of 4 by November 26, and yet Google did not even purport to finish the First Iterative Search until February 2022. Google did not provide any excuse for its delay aside from claiming that certain engineers were busy and/or on vacation.

283. Google has yet to complete the Second Iterative Search, particularly because Google delayed in producing data where it claimed to need publishers' consent. Google also has not produced full data from the "maybe_chrome_incognito" logs.

284. Fifth, Google undermined the Special Master process by first neglecting to inform the Plaintiffs about a set of preserved GAIA-keyed data for the named Plaintiffs. This became apparent during a Special Master live demo session when preserved data for the Calhoun Plaintiffs was discussed. Second, Google neglected to preserve Biscotti-keyed data associated with the Plaintiffs' GAIA-keyed data, some of which were preserved. I reviewed the production of preserved GAIA-keyed data (GOOG-BRWN-00847947 and GOOG-BRWN-00847948), and found that it included encrypted Biscotti identifiers marked with "encrypted_userid_DO_NOT_ACCESS_WITHOUT_PRIVACY_REVIEW {….

[159] 2022-03-25 Brown v. Google - UserAgent_IP UMA Search Script - AEO.pdf and 2022-03-17 Brown v. Google - UserAgent_IP Search Script – AEO.pdf

encryption_reasons: ENCRYPTION_REASON_BISCOTTI_JOINABILITY}." Yet Google represented that it did not search for and preserve any data tied to these Biscotti identifiers. That data would have been relevant for my analysis, but I was unable to review that data because Google did not preserve the data nor the Biscotti encryption key.

H.3 Class Identification

285. Throughout the class period, Google collected private browsing information from class members when they were not signed-into their Google accounts, and there are several ways to identify members of both Class I (Chrome class) and Class II (Non-Chrome class).

286. Google did not provide full access to its logs and data, with a Special Master process focused on data for the named Plaintiffs, so I was not able to at this point individually identify members of the two Classes – but that could be readily done in various ways using the data that Google stores and possibly in combination with some form of notification to potential class members.

H.3.1 Class I (Chrome Class)

287. As I will explain in the following paragraphs Google's internal logs and data repositories can be queried for purposes of identifying class members.

288. As described in more detail in Appendix G, Google has developed various server-side signals to detect Incognito traffic within its logs. For example, since at least July 4, 2021, Google has been logging a "maybe_chrome_incognito" field in [redacted] Google logs.[160] This field is a Boolean field, meaning that the value will always be either "true" or "false" (Liu Tr. 20:5-8). Google engineer Mandy Liu, the engineer principally responsible for developing the "maybe_chrome_incognito" field, testified that it would be possible for a Google employee with

[160] STIPULATION REGARDING GOOGLE'S "MAYBE_CHROME_INCOGNITO" FIELD

certain permissions and access to these logs to run a query to find "every single log entry where maybe_chrome_incognito shows up as a true" (Liu Tr. 41:25-42:12, 15:2-8). Based on my analysis of the data at issue, I agree with Ms. Liu that such a query is possible. The [redacted] fields where the "maybe_chrome_incognito" bit is being logged include each of Search ads, Display ads, and GAIA ads logs.[161] Accordingly, to find instances of users browsing within Incognito on non-Google website while signed out of Google, the query for maybe_chrome_incognito = "true" could be run on Google's Display ads logs. Moreover, Ms. Liu built a monitoring dashboard, which "uses" the "maybe_chrome_incognito" field and can thus be likewise used to query and isolate Incognito traffic.[162]

289. It is also my opinion that this query, whether through the logs or by way of Ms. Liu's dashboard, would accurately isolate traffic from Chrome Incognito users as opposed to non-Incognito users. As described in more detail in Appendix G, Google engineers drastically improved the accuracy of the "maybe_chrome_incognito" field over time, eventually reaching a point in February 2022 when data from the field matched well with Google Chrome's UMA "ground truth" metrics for Incognito usage.[163]

290. Specifically, from 2020 and onward, Google engineers working on detecting Incognito traffic within Google server-side logs, as well as the "maybe_chrome_incognito" field, were hoping to reach a point where the server-side detection logic would result in data showing that [redacted] of traffic was coming from Incognito, which was the figure shown to be the "ground truth"

[161] 2022-03-01 Brown v. Google - Google Letter re [redacted] Monitoring Dashboard.pdf

[162] Stipulation Regarding Google's 'Maybe_Chrome_Incognito' Field; Liu Tr. 36:16-37: 4; GOOG-CABR-05876933 (screenshot from the third-party cookie blocking dashboard); Liu Tr. 45:5-18).

[163] *See* GOOG-CABR-03849022 at -022 (Google engineer Bert Leung says: "we do have more updated stats from UMA…I believe we can use it as the 'ground truth' for [Incognito detection] analysis").

via UMA data (GOOG-CABR-04795991 at -996). Ms. Liu and Google achieved that goal in February 2022 after eliminating webview traffic from mobile apps[164] (Liu Tr. 40:3-20). Thereafter, Engineer Ms. Liu reported to her manager, Bert Leung, "Good news: Chrome incognito rate is [redacted]!" (GOOG-BRWN-00846508 at -508). Thus, Google engineers achieved extremely high accuracy with their Incognito detection logic. I describe in Appendix H.11 and Appendix I that the limited "maybe_chrome_incognito" production from Google demonstrates that the field accurately identifies Incognito data from the Second Iterative Search in the Special Masters process.

291. As this Incognito detection logic relies on X-Client-Data and user-agent that are readily available in Google's ads logs, the same logic could be applied retroactively to Google's ads logs containing X-Client-Data and user-agent (such as the logs Google engineers reviewed and analyzed in their log-based Incognito detection study in 2020[165]) to identify class members. I describe in Appendix H.11 and Appendix I that applying Google's Incognito detection methodology using X-Client-Data and user agent accurately identifies Incognito data from the Second Iterative Search in the Special Masters process.

292. The "maybe_chrome_incognito" field is not the only field that Google engineers have developed to detect Incognito traffic within Google server-side logs that rely on the absence of the X-Client-Data header. Others include "is_chrome_incognito" and "is_chrome_non_incognito."[166] As noted above, Google engineer Dr. Sadowski testified about [redacted] logs that contain the

[164] Webview are embedded Chrome browser instances in mobile apps and are not a part of this case.

[165] Google's Response to Interrogatory No. 34 and March 17, 2022 Supplemental Response to Interrogatory No. 34

[166] See e.g., Final Sadowski Reference Sheet listing several [redacted] logs containing is_chrome_incognito and is_chrome_non_incognito fields.

"is_chrome_non_incognito" field and two such logs that contain the "is_chrome_incognito" field.[167] These Incognito fields have been in use by Google in [redacted] logs since 2017 and are still in active use (Sadowski Tr. 70:11-18). As discussed earlier, on April 1, 2022, Google provided the proto comment for these fields. For the "is_chrome_incognito" field, the proto comment states [redacted] and for the "is_chrome_non_incognito_field", the proto comments states [redacted][168]

293. It is possible to use the fields identified above and Google dashboards to find event-level traffic attributable to Chrome Incognito, and those log entries can then be linked to people (including members of the Chrome class). I explained above in the Joinability section (Section F of this report) how various datapoints from Google log entries can be used to attribute particular log entries to people. For example, the combination of IP address and User Agent, or signed-in identifiers on non-Google websites (e.g., PPID and Analytics User ID), can be used to identify users. The [redacted] logs that contain the "maybe_chrome_incognito" field as well as the [redacted] logs containing "incognito" detection fields contain identifiers that can be used to find logs containing the same identifiers that track and store the user's IP address and User Agent. Accordingly, for any Google log entry where maybe_chrome_incognito is equal to "true," the combination of identifiers and the IP address and user agent for that log entry can be used to identify class members. Moreover, as noted above, the IP addresses and User Agent strings can be used to join a user's private browsing activities on non-Google websites with the user's Google

[167] Final Sadowski Reference Sheet.pdf

[168] 2022-04-01 Brown v. Google - Google Letter.pdf

account identity, and Google could then notify the class member via the email address associated with that Google account.[169]

294. To be sure, this method I have outlined will only work for the whole class period insofar as Google retained the necessary data. But I understand that Google is routinely deleting data from many of its server-side logs, including from the [redacted] logs that contain the "maybe_chrome_incognito" field. Google could have preserved a set of data for the purpose of identifying class members in this litigation, but Google did not.

H.3.2 All Class Members

295. There are other ways to identify class members. Because both Classes are limited to Google Account holders, they are necessarily limited to individuals where Google currently has (or at least at some point had) contact information. Google both before and throughout the class period collected information from individuals when they signed up for a Google Account. Also, Google Account holders have an associated email address, where Google will have records of that email and other contact information.

296. Given that Google should have email addresses for all or almost all class members (the exception being people who canceled their Google Account and Google therefore deleted that record), one easy way to identify class members would be by sending emails to those email addresses. There are various ways to limit that email notification to individuals who are likely members of the two Classes. As one initial step, notification could be limited to email accounts associated with people in the United States, based on Google's own records. Based on my analysis

[169] Plaintiffs can obtain their IP address and user agent from their devices, GAIA/Biscotti/Zwieback logs and UMA logs as discussed earlier. Plaintiffs can also obtain their IP address and user agent from Google Take Out. For example, **Exhibit D** contains Plaintiffs' Google Take Out files containing their Google Account information, IP address and user agent.

in this case, given the broad usage of the private browsing modes throughout the class period, it is my opinion that most of those people would be class members. If required, it would also be possible to use Google data to further limit email notification to accounts associated with some private browsing behavior.

297. In addition to email notification, or as an alternative, for Class I, Google could provide notification to Chrome Incognito users by way of the Incognito screen. For example, if someone started using Chrome Incognito, the screen that Google shows could be updated to provide notification and a link to additional information.

298. Alternatively, Google could require non-Google websites to provide pop-up notifications whenever users visit websites for which Google will collect private browsing information, providing another way for Class Members to receive notification.

299. I understand that Counsel has retained a class notification expert who will offer additional opinions regarding other forms of notification, including publication.

300. After notification, there are various ways to assess the extent to which Google collected private browsing information during the class period such that the person would be a Class Member. For example, individuals could provide additional information regarding their private browsing during the class period, or those users could provide certain identifying information to be used to search for records in Google's data sources. Google's data could readily be used to verify whether that user was a Google Account holder during the class period and otherwise assess any response to any notifications.

301. Both before and throughout the class period, Google assigned every Google account holder a unique GAIA ID associated with their email address. Through the Special Master process, Google produced the GAIA IDs of all the Plaintiffs. For example, GOOG-BRWN-00819269

shows the GAIA ID associated with Plaintiff Chasom Brown's email address. Google further stores the date and time of Google account creation, IP address at the time of account creation, as well as IP address, user agent, and date and time of Google account log-in and log-out activities in both private browsing and non-private browsing modes (*e.g.*, GOOG-BRWN-00817563, GOOG-BRWN-00819272, GOOG-BRWN-00819441, GOOG-BRWN-00819830, GOOG-BRWN-00820370).

302. Class members could provide additional information about their private browsing history, such as by identifying non-Google websites they visited during the class period and any available IP address and user agent information. If the IP address and user agent are not available, this information may be obtained from Google's own GAIA records as discussed above. The IP address and user agent in the GAIA records would be the same IP address and user agent when the user is signed-out of their Google account while connected to the same network. Signing into or out of a Google Account does not change a device's IP address and user agent.

303. It would also be possible to search Google's signed-out logs to verify the existence of records corresponding to the user and establishing membership in one or both Classes. Every one of the located records would be from browsing activities on websites containing at least one Google tracking beacon – since that's how Google intercepted the data contained in those logs in the first place.

304. If a user has signed into their Google Account in private browsing mode within the [redacted], then Google can also retrieve the user's Biscotti IDs from GAIA logs and use those Biscotti IDs to locate the user's private browsing records while they are not signed into Google from the Biscotti logs.

305. In addition, from any private browsing session (Incognito in Chrome or private browsing mode in other browsers), a user can (with the help of technology professionals) extract cookies from their browsers in private browsing mode (e.g., Biscotti cookies, GFP cookies and _ga cookies containing Analytics CIDs). These cookies are unique to that private browsing session and can be used to identify the user's browsing activities in that session. Upon receiving these cookies, it is possible to locate the user's private browsing data in Google logs and identify every website visited during that session that Google tracks.

306. Furthermore, a user can extract their non-Google sign-in IDs (with the help of technology professionals) from websites that they visit in private browsing mode, including PPID, Analytics User ID and Google Ads User ID. Upon receiving these IDs, Google can locate the user's private browsing data from those websites in logs.

307. I provide a more in-depth discussion of the different ways to identify the Chrome Incognito class, including Google's own Incognito detection methods, in Appendix G.

I. Attempt: Google's attempted interception and collection uniformly impacted all class members

308. Even without Google preserving or providing data on specific users and specific interceptions, Google throughout the class period and for all class members at a minimum attempted to intercept and collect data from those private browsing communications. Google designed its tracking and advertising code to be embedded on any website and to be agnostic to the specific browser and device for web browsing (Berntson March 18, 2022 Tr. 141:5-142:15).

309. Google's tracking and advertising code are embedded on such a large percentage of websites that it is effectively impossible for any user browsing in private browsing mode – including all class members in this action – to not encounter at least one Google tracking or advertising code while browsing privately. My analysis in this case confirms that these Google

tracking and advertising code would have functioned in a way that at a minimum attempted to intercept and obtain information from the private browsing communications at issue in this lawsuit when people visited non-Google websites.

310. Several Google and third-party studies show that Google tracking beacons are prevalent on the Internet. A 2020 report by DuckDuckGo shows that "Google-owned trackers are on over 85% of the top 50K sites".[170] From DuckDuckGo's survey of 29,578 websites, Google's display ad "doubleclick.net is being used on sites not owned by Google ~98% of the time."[171] Another study by Ghostery shows that in 2017 and 2020, Google was on ~64% and ~86% of US websites respectively.[172] Yet another study by Princeton University in 2016 shows that Google trackers are on more than 70% of the top million websites.[173] Google's internal documents show that Analytics tracking beacons alone are installed on [redacted] of the top [redacted] websites throughout the class period—to say nothing of the many other tracking beacons at issue, such as Ad Manager AdSense, and conversion tracking beacons.[174]

311. Estimating [redacted] of websites visited in private browsing mode use at least one Google tracking beacon, then the probability of hitting at least one Google tracking website as a function, *F,* of the number of websites visited, *n*, is $F(n) = 1 - (1 - 0.7)^n$. Thus, if a user visits just 2 non-

[170] https://spreadprivacy.com/duckduckgo-tracker-radar/ (Last accessed on April 11, 2022).

[171] https://spreadprivacy.com/duckduckgo-tracker-radar/ (Last accessed on April 11, 2022).

[172] https://www.ghostery.com/blog/tracking-the-trackers-2020-web-trackings-opaque-business-model-of-selling-users (Last accessed on April 11, 2022).

[173] https://www.cs.princeton.edu/~arvindn/publications/OpenWPM_1_million_site_tracking_measurement.pdf (Last accessed on April 11, 2022).

[174] GOOG-BRWN-00490767 at -772 (Last updated Nov. 2020), GOOG-CABR-00381312 at -321 (Last updated Nov. 2019), GOOG-BRWN-00491711 at -718 (Last updated Mar. 2018), and GOOG-BRWN-00424253 at -295 (Last updated Sept. 2017).

Google websites, the probability of hitting at least one Google tracking beacon is [redacted]; 4 non-Google websites, the probability is [redacted]; and 10 websites, the probability is [redacted]. Google's own metric shows that the average number of Incognito sessions per user within just a [redacted] period is between [redacted] (GOOG-CABR-04648434 at -442). Even if a user visits just one non-Google website per week, over the course of the class period, it is virtually impossible to not come across at least one non-Google website with at least one Google tracking beacon.

312. Meanwhile, Google does not provide users with any tool to escape these Google tracking beacons on non-Google websites. The Plaintiffs in this case are alleging that Google portrayed private browsing mode, including Incognito mode, as the control to prevent Google from tracking them across non-Google websites. But Google tracking beacons nevertheless uniformly attempt to intercept users' private browsing information during their visits to non-Google websites, including while signed out of any Google account.

313. Google through its pervasive tracking amasses massive amounts of user data. One internal Google presentation explained that "Google is a data driven company" and that "Google's mission is to organize the world's information and make it universally accessible and useful" (GOOG-BRWN-00561172 at -75). That same presentation explained that Google "crossed [redacted] of data in 2017" and that "[redacted] of [this] data is User Data"—i.e., [redacted]. Assuming very conservatively that [redacted] of this user data is from users' private browsing sessions, Google was storing [redacted] of private browsing information. And just [redacted] is the equivalent to "about [redacted]."[175] All of the data Google collects and stores, including private browsing information collected from users' visits to non-Google websites

[175] https://adeptia.com/blog/surprising-things-you-dont-know-about-big-data (Last accessed on April 11, 2022).

while signed out of any Google account, contributes to the machine learning algorithms and modeling discussed above, which Google relies on to market itself to advertisers and generate revenue.

314. Google's Supplemental Response to Interrogatory No. 5 lists "the 25 top-level domains with the highest website traffic, as measured by total Ad Manager Ad Requests, for Ad Manager publishers with a U.S. time zone and billing address for June 2, 2020":

accuweather.com
apple.com
britannica.com
buzzfeed.com
cbslocal.com
chess.com
cnn.com
dailymail.co.uk
ebay.com
fandom.com
foxnews.com
kohls.com
nypost.com
nytimes.com
pch.com
quizlet.com
realtor.com
reddit.com
target.com
usatoday.com
washingtonpost.com
weather.com
webmd.com
youtube.com
zynga.com

315. Google further lists, for Google Analytics "the 25 top-level domains registered in US time zones with the highest traffic, as measured by visits from unique Client IDs, for June 2, 2020":

accuweather.com
businessnsider.com
change.org
chaturbate.com
condenast.com
condenastinternational.com
developer.yummly.com
grammarly.com
hearst.com
howto.gov
indeed.com
linkedin.com
mobile.nytimes.com
nexstar.tv
nytimes.com
paypal.com
pornhub.com
privy.com
roblox.com
surveymonkey.com
townnews.com
vdo.ai
washingtonpost.com
wikia.com
worldometers.info

316. Google's internal document GOOG-CABR-05280050 titled "Publishers by Chrome Incognito %" lists [redacted] US publishers according to the [redacted]" field. This document shows that Google keeps detailed Incognito usage information broken down by country and publishers. All these websites implement at least one Google tracking beacon.

317. Google's MARCH 17, 2022 Supplemental Response to Interrogatory No. 33 lists [redacted] Google Analytics Accounts and [redacted] Google Ad Manager Publisher IDs with at least one query and a billing address in the United States. Each of these entities may be associated with more than one website (GOOG-CABR-04294727 states "many publishers have 50+ sites"). All these websites implement at least one Google tracking beacon.

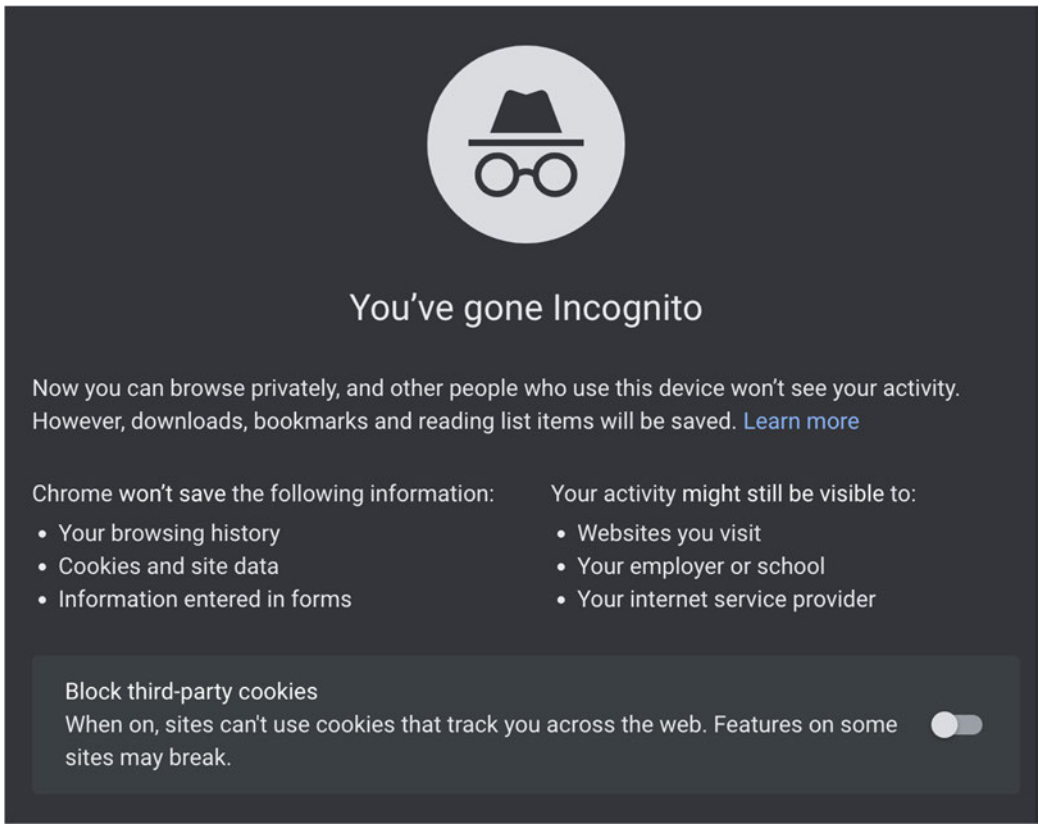
J. Incognito Functionality: Throughout the class period, Google's Chrome Incognito mode functioned in ways that were different than represented

318. Throughout the class period, Google's Chrome Incognito mode functioned in ways that differed from how Google represented it would function.

J.1 The Incognito Splash Screen

319. Currently, when a user opens a Chrome Incognito window, they are automatically shown the Incognito Splash Screen below. On iOS devices, the "Block third-party cookies" section (i.e., [redacted]) is not shown since [redacted] was never launched on iOS devices (GOOG-CABR-04986312). The [redacted] switch is also not shown prior to May 2020 on all devices (GOOG-CABR-04820567 at -577). Previous versions of the Incognito splash screen message can be found in GOOG-BRWN-00555223. Google has admitted in this case that "since June 1, 2016, a version of the full-page Incognito Notice (which Plaintiffs refer to as Incognito Splash Screen) is designed to be shown to a user every time a user chooses to enter Incognito mode using the Chrome browser. Based on a reasonable investigation conducted by Google to date, Google admits that, since June 1, 2016, a version of the full-page Incognito Notice appeared to every user each

time they enabled Google's Incognito mode" (Google's Response to RFA No. 26). Likewise, Google's internal documents state that "we show the user every time they open the Incognito tab about what they're guaranteed" (GOOG-CABR-00429070 at 186).



320. The Incognito Splash Screen represents that "Chrome won't save the following information: Your browsing history, cookies and site data, information entered in forms." This statement is false as a technical matter. In fact, Chrome does save browsing history as well as cookies and site data within Incognito sessions—at a minimum for the duration of the Incognito session. This is significant in part because that is part of how Google intercepts and collects private browsing information – by Chrome collecting, saving and then delivering that information to Google's servers, where Google then stores and exploits that information. As former Google engineer Rory McClelland explained, "I would argue that we do store it, just in memory only. One of the main misunderstanding we have with IM [Incognito Mode] is with people not closing their

sessions, which would be reasonable if they believed we never store the data…it does have impact on the accumulation of cookies (and sign-ins) that allow the user to be tracked, even in Incognito mode" (GOOG-BRWN-00699213 at -214). While the Chrome browser itself may clear certain browsing data after a user closes all Incognito windows, Chrome certainly sends browsing information, cookies, and site data to Google servers while the user is browsing within the Incognito session. These data are logged in various logs and storage as I have described in Section D of this report.

321. In addition, since June 1, 2016, Google has consistently identified "websites you visit", "employer" and "internet service provider" as entities to whom browsing information may be visible. Conspicuously, Google itself is left out from the list. "Google admits that, since June 1, 2016, the full-page Incognito Notice (which Plaintiffs refer to as Incognito Splash Screen) has not specifically identified Google by name as an entity that 'can view a user's activity in private browsing mode" (Google's Response to RFA No. 27). Meanwhile, Google's internal documentation admits "Chrome Incognito mode: … does **not** provide anonymity/invisibility towards any other party (such as the websites you visit, or Google)" (GOOG-BRWN-00390418 at -418).

322. Google's internal documents recognize the existence of "common misconceptions about private mode" including that it "hides browsing activity from Google" - noting "Participants don't want Google to collect data in Incognito, especially data tied to their identity" and that "Users are concerned about (Google) collecting data in Incognito" and "Given that participants overestimate

private mode protections, participants are surprised and feel misled when made aware of private mode vulnerabilities" (GOOG-BRWN-00051239 at -267, -272, and -273).[176]

323. Google has been actively tracking Incognito usage, including through UMA, [redacted] location logs, and, more recently, through the logging of a maybe_chrome_incognito bit in various ads event logs as I have discussed earlier. As Google engineers agree, the availability of an Incognito tracking signal "violates the concept of Incognito" (GOOG-BRWN-00845585) and "is a PR nightmare waiting to happen" (GOOG-BRWN-00176481 in connection to discussion of the "extra header to identify Incognito sessions" in GOOG-BRWN-00176433).

324. The current version of the Incognito Splash Screen (shown above) contains a [redacted] toggle that by default blocks "third-party cookies." Google launched this feature in May 2020 as mentioned earlier. Google's representation that "sites can't use cookies that track you across the web" does not inform users that [redacted] does not block Google's first-party cookies set on non-Google websites. Given the broad coverage of Google's tracking beacons, most highly visited websites have Google's first-party cookies set[177]; thus, while [redacted] prevents some non-Google service providers from tracking users across the web, it does not prevent Google from tracking users across the web. Indeed, a Google Chrome engineer characterized the notion that Incognito prevents "ad tracking -- ad targeting and tracking" as "one of the highest misconceptions" (Halavati Depo. 76: 1-4 testifying in reference to GOOG-BRWN-00042388).

[176] *See also* Halavati Depo. 77: 17-21 (affirming that for participants in study referenced in GOOG-BRWN-00042388, the belief that incognito mode hides browsing activity from Google is one of the common misconceptions among users).

[177] E.g., GOOG-BRWN-00490767 at -772 showing Google Analytics' tracking beacons installed on [redacted] of the top [redacted] websites. These tracking beacons set and use first party cookies. And GOOG-CABR-04118321 at -336 showing first party cookies set by Google's Ad Manager and AdSense tracking beacons.

J.2 Google's Chrome Incognito mode functioned in ways that were different than how Google represented it

325. Google's internal documents acknowledge that "The Incognito Problem… Incognito Confuses People" (GOOG-BRWN-00140297 at -297 and -298), "Users value Incognito, but are confused by it…Incognito is widely known and appreciated, but misunderstood" (GOOG-BRWN-00166360 at -367) and "Private browsing has a problem. In both internal and external research, we've seen that the privacy properties of Incognito aren't well understood, to say the least" (GOOG-BRWN-00164626). In describing Incognito, Google often must perform "mental and linguistic gymnastics"[178]: Google CEO Sundar Pichai was told: "Do not use the words private, confidential, anonymous, off-the-record when describing benefits of Incognito mode; These words run the risk of exacerbating known misconceptions about protections Incognito mode provides" (GOOG-BRWN-00048967.C at -968.C). Mr. Pichai ultimately decided that he did not "want to put incognito under the spotlight" (GOOG-BRWN-00388293 at -293). Similarly, Google employees lament, "We are limited in how strongly we can market Incognito because it's not truly private, thus requiring really fuzzy, hedging language that is almost more damaging" (GOOG-BRWN-00406065 at -067).

326. These misconceptions start from the name Incognito itself and the spy guy icon. As Google engineer Christopher Palmer admitted, "the blame for people's misconceptions about Incognito Mode is due to that name and branding, as I have argued repeatedly" (GOOG-CABR-03958924 at -925). Another Google engineer says "we need to stop calling it Incognito and stop using a Spy Guy icon… I know the 'incognito' war was waged and lost years ago, but do you remember why?

[178] GOOG-CABR-04177585: "we have to do a bunch of mental and linguistic gymnastics when talking about incognito logging"

It has always been a misleading name" (GOOG-BRWN-00475063 at -065). Other Google engineers opine "regardless of the name, the icon should always have been http://simpsons.wikia.com/wiki/Guy_Incognito...which also accurately conveys the level of privacy it provides I think." And "They didn't believe me that people would get confused" (GOOG-BRWN-00475063 at -065).

327. Other Google documents state: "Yes, to begin at the beginning, I tried typing 'define incognito' into Google search… seems the name itself may be a cause of confusion: '1. (of a person) having one's true identity concealed'" (GOOG-BRWN-00061607) and "common misconceptions exist … Research tells us the current Incognito icon is not alleviating the misconceptions" (GOOG-BRWN-00072055 at -070 and 071). Studies have also shown that "many users believe private browsing gives them anonymity", "many users believe private browsing obscures location", "many users believe private browsing actively blocks ad tracking", "many users believe private browsing prevents search engines from remembering their searches", "most users believe that their private browsing activity won't be remembered by Google, even after they sign into their account" (GOOG-BRWN-00061607), and "Users believe that they can protect

themselves against bad actors or hackers by logging-in to sensitive accounts in Incognito" (GOOG-BRWN-00165706 at -712).

328. These Google-created user misconceptions lead users to "overestimate the privacy protections that Incognito/private browsing provide, potentially putting users at risk in the moments when they expect the most privacy. We've seen these misconceptions consistently in product research, academic research, and several [redacted] studies. From the brand perspective, Incognito mode is seen as 'private' but still might be overpromising (most users grasp the basics, but there is still confusion around what it does and expectations around being 'anonymous')" (GOOG-BRWN-00156752 at -788).

329. Google's internal documents also acknowledge: "The idea that Incognito can provide anonymity creates a false sense of security in returning Incognito users", that "Users may not do the things they need to do in order to remain safe online" (GOOG-CABR-04489649 at -655). For example, "The misconception that Incognito can obscure location can lead users to ineffectively avoid danger…E.g., Users who are fleeing abusive partners may not take the appropriate steps to turn-off geolocation settings because they believe that IM [Incognito Mode] is doing this already" (GOOG-CABR-05145417 at -421 to -422).

330. In addition to the branding message, Incognito's features are also sources of misconceptions. As Google acknowledges, "There is a gap between our promises about Chrome Incognito mode, their delivery, and user expectations. This is partly because of not being able to transfer the message to the users, and partly due to missing/incomplete features" (GOOG-BRWN-00047399). Google engineers expressed what they wished "were different about Incognito -More privacy, -It was more clear on what it does/doesn't do Google and others are still tracking me – it's not truly private…Each tab a segregated sandbox…what happens in an Incognito tab is trapped

in an Incognito tab (Incognito cookies can't be read across Incognito tabs) -It was easier to understand…-That is were anonymized…-I wish Incognito fulfilled the privacy expectations…That it protected my data from Google and other websites" (GOOG-BRWN-00148738 at -745 and -746). As Incognito is designed, Google certainly tracks users and does not protect user data from Google.

331. To the point about Incognito tabs, Google did not implement each Incognito tab as a "segregated sandbox" such that "what happens in an Incognito tab is trapped in an Incognito tab" and Incognito cookies are not shared across different tabs (GOOG-BRWN-00028052 at -072). Instead, Chrome implements a shared cookie jar across all Incognito windows and tabs (referred to as session-based tracking).[179] As a "user continues using Incognito, websites can start writing cookies and store different sort of information about the user in the browser's storage, and these can be used for further tracking while user has not still closed the Incognito session." As a result, if a user goes to a website "in Incognito mode and continue[s the] Incognito session and go[es] to that website again, they may know that [the user is] the previous person and continue serving to [the user] as the previous person" (Halavati Depo. 78: 3-14). This means a user needs to close every Incognito window to end the Incognito session and clear the cookies and browsing history accumulated during the session.[180]

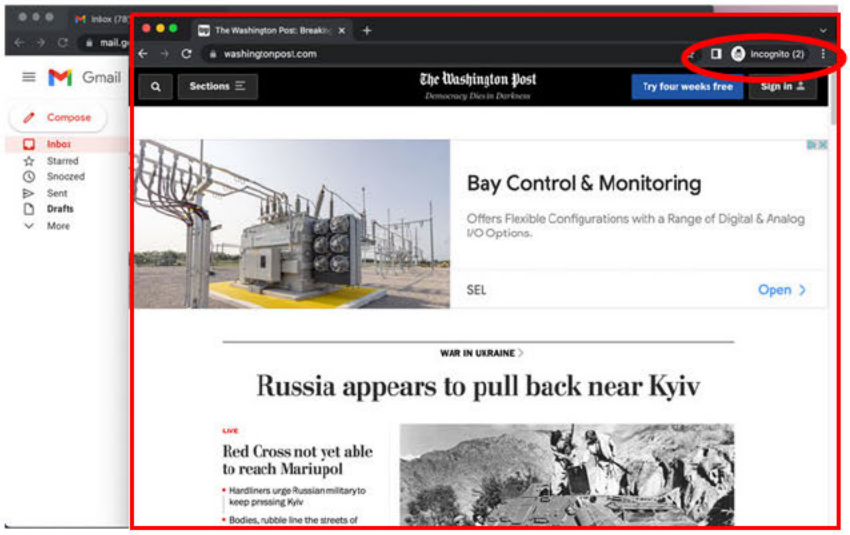
[179] GOOG-BRWN-00605636 at -636: "Closing Incognito will clear all cookies and site info related to the session Many users are unaware of session based tracking" and Halavati Depo. 113: 10-17 (recognizing differences between Chrome and Safari in that, in Chrome, Google "consider[s] all tabs in one window in a common session, but in Safari, they consider each tab a separate, isolated session. . . Yes. It can result in more privacy [in reference to question about tracking protection like Safari's], but less functionality" (citing GOOG-CABR-00545997)).

[180] *See* Halavati Depo. 81: 1-3 (confirming understanding that from August 2016 until the addition of third-party cookie blocking in Incognito mode, incognito mode did not prevent session-based tracking including through the use of third-party cookies).

332. Google's internal document shows that "Most users are not aware of session-based tracking", "Majority of user [sic] expect no session-based tracking" and "Many users don't understand and are not aware of session-based tracking" (GOOG-BRWN-00051239 at -282). A Google engineer wrote "another common issue we encounter is that users don't recognize that closing an Incognito browser window is not sufficient to clear the session tokens…*all* Incognito browser windows must be closed, even minimized or background windows the user hasn't interacted with" (GOOG-BRWN-00410821at -826). This is a significant problem, because "There may already be Incognito…windows open that the user doesn't see, which could have unintended consequences for the user if they end up sharing state with an existing Incognito…session" (GOOG-BRWN-00410821 at -824). For example, "a previous user may have left one [window] open and hidden, possibly getting the new user to leak data into the previous user's accounts without realizing it" (GOOG-BRWN-00410821at -825). While Chrome Guest mode has an "'Exit Guest' button in the user menu (chip in the top-right corner) which closes all Guest windows" (GOOG-BRWN-00410821 at -826), Chrome Incognito mode does not have a button for users to close all Incognito windows. Furthermore, "The more time [a user spends] in an Incognito session, the more cookies and trackers are accumulated, which can make the experience less private … Anonymity decays as you spend more time in Incognito ← hard to explain without freaking people out" (GOOG-BRWN-00571757 at -758 and -761).

333. Another aspect of Incognito is Google's tracking behavior when a user is signed into their Google account in Incognito mode. "Many users believe private browsing gives them anonymity . . . prevents Google from seeing their activity, even after they sign into their account" (GOOG-BRWN-00186446 at -448). Apparently, this is unknown even for Google engineers, "Incognito on google.com… Users don't know how Incognito works when they sign in to google.com. Neither

do we!" (GOOG-BRWN-00166360 at -393). Unlike Chrome regular mode where a user's account icon shows up on every tab and window when signed-in, in Incognito mode, when a user signs into their Gmail in one tab or window, for example, other Incognito tabs and windows do not show the user's Google account icon. Instead, the user sees an Incognito icon on all other tabs and windows. See screen capture below of two Incognito windows, one overlaid on top of another. In this example, the user is signed-into Google in the bottom Incognito window but the top Incognito window has no indication of this sign-in activity in the other window. It would not be unreasonable to consider the top window to be not signed-into Google; however, Google deems browsing activities on all tabs and windows where the user did not sign into Google to be "authenticated" and stores private communication on those tabs and windows in GAIA logs.



334. As Google's own study reveals, there is a fundamental "Misalignment between back-end functionality and user's mental models of Incognito mode" (GOOG-CABR-05163088 at -091). In an internal presentation, a Google employee used the following image to summarize Incognito mode (GOOG-CABR-05756666 at -684):

K. Incognito Changes: Google could have changed Chrome Incognito mode to match user expectations and address misconceptions, but Google chose not to

335. Google engineers appear to have understood user expectations around Incognito and called for Google to honor that expectation: "When a user asks for Incognito, it's a huge gift, it's a vulnerable moment, the user is telling us I feel at risk. We need to make a better commitment in respecting that intent" (GOOG-BRWN-00466897). In internal Google documentations, Google engineers also appear to frequently express exasperation about the misleading nature of Incognito. The question raised by Google engineers is this "I know there's sometimes a mismatch between actual/expected behavior when users are in Incognito mode. Are we trying to realign the mismatch by making incognito more…functionally incognito? Or by rebranding to make it clear that Incognito doesn't promise too much?" (GOOG-CABR-05741188 at -189). Along these two veins – Incognito functionality and promise, Google engineers have proposed many solutions, which would be viable in my opinion, to no avail.

336. One Google document states the Incognito misconceptions "could give users the false impression of privacy in a moment when they expect it most, leading them to inadvertently share sensitive personal data. We therefore must be bolder about saying *who sees what* as we browse the web, particularly during these moments. This should begin with creative changes to naming conventions, iconography, explanation strings, and/or the function of Incognito itself" (GOOG-CABR-04466637). Around this "bolder" "who sees what" issue, Google could have made it clear to users that Incognito is not protecting users from Google as illustrated by Google engineers (GOOG-BRWN-00048773). I agree with this analysis by Google's own engineers.

337. Google rejected this iconography proposal, McClelland Tr. at 58:19-59:6, even though Google knew there was a user "misconception" that "private browsing mode would protect" users from Google, *id.* at 45:24-46:10, and that "[w]here there is, in any product, a gap between what actually happens and the user's mental model, there is a risk that the user makes a misinformed decision around the use of their data with privacy." *Id.* at 46:11–47:13.

338. Other engineers suggest "Rachet Down Incognito" – "Use less loaded terms and iconography (e.g. 'Temporary Mode', Recycle Icon, or whatever). Simplify the disclosure/disclaimer as a result. Benefit: No longer over-promising and under-delivering" (GOOG-BRWN-00140297 at -315). Some recommend that Google fix "real problems" like "publishing exactly what data we normally collect and how we will use it, now and forever (which we do NOT sufficiently do IMO) …" (GOOG-BRWN-00226894). Some recommend "extend InCognito mode so that it is high privacy by default (all the privacy options are basically turned on). This would really be the easiest way to provide users with a mode that instantly provides them with complete privacy!" (GOOG-BRWN-00408322). I agree with this analysis by Google's own engineers.

339. Still other Google engineers proposed numerous feature changes. In one document, Google employees recognized that changes including disabling all cookies, disabling pixels, detecting and warning users of sign-in actions or when their personally identifying information is entered, hiding IP addresses, and disabling location sharing, could more closely align Incognito features with user expectations (GOOG-BRWN-00152199 at -203). Another document suggests "Set a default cookie expiration when cookies are set in Incognito…Remove user agent strings from logs; Remove any of the data that is use[d] for fingerprinting…Warn users that logging in makes their identity known on both Google and any other site…Fuzz data collection (timestamps, etc) when Incognito" (GOOG-BRWN-00157528 at -528). Other feature changes included "Incognito: Users should have the ability to opt out of personalized experiences and data sharing, and can do so without any difficulty (that is much more effective than today)" (GOOG-CABR-05280966 at -022), "reduce fingerprinting vectors to the point where few people are individually distinguishable" (GOOG-BRWN-00475063 at -064), ""mask[] IP address" to "deliver stronger

guarantees" (GOOG-BRWN-00650016), "remove the IP address, the user agent, and the cookie ID from the logs after the Incognito session is over…I really like the idea that we eventually don't know that a given log entry was from an Incognito session" (GOOG-BRWN-00615433 at -434), "We could consider dropping all incognito-mode queries from [redacted], or, later, not including them in [redacted]…when a user goes incognito, we could alto-request a deletion directly after every search is done" (GOOG-BRWN-00615433 at -435), and "We are looking at a proposal where we stop logging data from Incognito sessions" (GOOG-BRWN-00615433 at -435). In my opinion these approaches would be helpful and viable.

340. One Google document titled "a few thoughtful ideas" suggests "Create a **true** incognito mechanism – no tracking, no cookies, nada", "Start all new Chrome tabs in incognito mode. Make tracking opt in, not opt out", and "Make it easy and simple for people to take-out their data" (GOOG-BRWN-00847949 at -951). The custodian of this document is Lorraine Twohill, Google's Chief Marketing Officer. This approach matches my understanding of what online privacy means, which Google never provided.

341. Google was aware of the issues with Incognito long before the class period; however, Google did not make changes suggested by its engineers. Even after this lawsuit began, Google continued to track users in private browsing modes and used private communication data to fuel its [redacted] advertisement business. The obvious inference I can make is that Google values its advertising business more than the privacy of its [redacted] of users in the United States.

L. Google can change its processes going forward and purge its systems of private browsing information

342. As discussed earlier, there are various changes that Google could make that would allow for privacy going forward, with Google no longer collecting, storing, and using private browsing mode information.

343. As noted above, Google has begun logging event-level data as Incognito traffic in at least [redacted] Google logs ([redacted] logs with the "maybe_chrome_incognito" field, [redacted] logs with the "is_chrome_non_incognito" field, and [redacted] logs with the "is_chrome_incognito" field). I understand that Google has not clarified one way or the other whether there are any other fields (aside from X-Client-Data) in Google logs that track "incognito" traffic. Also as noted above, Google could run a query to locate event-level data where these fields show up as either "true" or "false." For Incognito events, Google can identify the associated pseudonymous identifiers on non-Google websites and thus delete from its systems data that it should have never collected in the first place. Even if Google is sometimes unable to distinguish between private browsing data and non-private browsing data, they still have the ability to delete all private browsing data by deleting any data that might be private browsing data.

344. In addition, traces of users' private browsing information (from Incognito mode and other browsers) are not limited to event-level traffic within Google logs. Google has also used private browsing information to build and improve Google's algorithms and models, which Google continues to use. To prevent further benefits from Google's past collection of private browsing information, these Google algorithms and models would need to be deleted and rebuilt—this time without any private browsing information.

345. For example, to effectuate personalized advertising, Google's machine learning algorithms use various signals to predict users' future interests (GOOG-CABR-04616196 at -221 to -223).

Machine learning algorithms are also used extensively for serving advertisements. Relevant machine learning tools include Sibyl for ad ranking, Laser for personalization, SmartASS for ad selection, among others (GOOG-BRWN-00535100 at -156). These tools have benefited from and rely upon all browsing information, including private browsing information.

346. It is not possible at this point to "unring the bell" with these tools. Instead, these tools would need to be deleted and rebuilt or retrained, otherwise Google would still be benefiting from the private browsing data that it collected.

347. As another example, also as discussed above, Google relies on machine learning algorithms to model and predict conversions, and these algorithms likewise rely on both private and non-private browsing information. One specific tool is Google's [redacted] program, through which Google uses Google's large volume of "observed conversions" to build models for Google to estimate cross-device conversions when direct conversion tracking cannot be observed (*See*, GOOG-CABR-03670580). GOOG-CABR-03669893 at -900 shows simulated or modeled conversions are based in part on "unattributed conversions", which may include private browsing information. It would not be possible for Google to strip out from [redacted] or any similar tool only the private browsing information that has been used to build and perfect it. Instead, to the extent there is a need to address how Google continues to benefit from the conduct at issue with [redacted] and similar tools and algorithms, those tools and algorithms would need to be deleted and rebuilt from scratch—this time without the benefit of private browsing information.

348. These examples are not exhaustive. I understand that Counsel sought discovery from Google concerning the full scope of Google's "use of information collected from users within a private browsing mode for purposes of improving and developing Google services, products, and algorithms" (Dkt. 411-1 at 12). But Google objected, claiming that "This topic is overly broad and

amorphous and would implicate dozens of Google business units and products" (*Id.*). Although I have not been provided with information about these "dozens of Google business units and products," this Google statement does not surprise me. As illustrated by the above examples, Google uses the information it collects from users' browsing sessions, private or regular, to improve and develop Google products and algorithms.

349. I have not formulated any opinions regarding what, if any, changes Google might make to any disclosures in connection with any injunctive relief.

IX. CONCLUSION & RIGHT TO SUPPLEMENT

350. My investigation is ongoing. I reserve the right to update or supplement this report in reliance upon whatever further information becomes available.

Respectfully submitted by,

/s/ Jonathan E. Hochman
JE Hochman & Associates LLC
Date: April 15, 2022

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

APPENDIX A

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

APPENDIX B

Sealed Entirely

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

APPENDIX C

Produced Electronically

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

APPENDIX D

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

APPENDIX E

Produced Electronically

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

APPENDIX F

Appendix G Chrome Class Identification

Content

1. There are a number of ways to identify members of Class I based on the information collected and stored by Google, and Google employees have admitted as much in internal discussions. Throughout the class period, Google has detected, monitored, and logged Incognito browsing activities through two primary means: (1) UMA; and (2) x-client-data in combination with user agent. Google has implemented several internal "Dashboards" to monitor and display Incognito browsing traffic, both in terms of individual events and in terms of aggregate statistics. These Google Dashboards, as well as the logic used to implement Incognito detection on Google's server side, can be used to identify the Chrome Incognito class (Class I). I will describe these methods and the Dashboards in this Appendix.

APPENDIX G.1 USER METRICS ANALYSIS (UMA)

2. It is my opinion that Google can use its records of UMA data to identify members of Class I. This class identification method will not work for all members of Class I because not every Chrome instance is selected by Google to upload UMA data.[1] And this method is being impaired by Google's ongoing deletion of private browsing information. The data is only retained for [redacted].[2]

[1] GOOG-CABR-04477378 at -381: "The metrics team controls how many of the consenting users actually upload data to Google."
[2] GOOG-CABR-04801490

3. As discussed in Section F of the main report, UMA is a built-in feature in Chrome that reports extremely detailed Chrome usage information to Google, including Incognito mode usage. UMA has been in existence throughout the class period, with Google obtaining information regarding Incognito mode usage throughout the class period. Users have no ability to delete their sensitive data stored in Google's UMA logs (raw or aggregated).[3]

4. Each Chrome instance or install is associated with a unique identifier called the UMA ID or [redacted]. The UMA [redacted] identifier is a stable identifier that is created when a user installs a Chrome instance and does not opt-out of UMA data collection.[4] A user can obtain their UMA ID ([redacted]) by typing [redacted] in a Chrome browser URL box and copying the [redacted] value under [redacted]. This ID can be retrieved for each device that a user owns. When multiple user accounts exist on a device, each user can retrieve their own UMA ID by signing into their own device account.

5. For each UMA ID, Google can look up the raw UMA data periodically uploaded to Google from the user's device in a UMA log called the [redacted]. This log [redacted][5] This raw UMA data log is one of the key Google logs for identifying Incognito usage on an individual user basis.

[3] GOOG-BRWN-00032906 and GOOG-CABR-00484121at -121: "The retention policy for the received data is 1.5 years. But we do keep aggregates of it indefinitely … The existing collected data cannot be deleted by the user"
[4] There are a few limited scenarios when UMA ID is reset, including when a user factory resets the device or toggles off the consent for UMA data collection. For the vast majority of the use cases, the UMA ID does not change. *See e.g.*, GOOG-CABR-04486185 at -186: "There are multiple reasons why UMA client ID can be reset, one of which includes the device being powerwashed (also known as a factory reset and switch to Dev mode). In these instances, new UMA client ID is created after a device is powerwashed…" and Footnote 3: "UMA Client IDs can also be reset by entering Dev mode, switching channels (i.e. from Stable to Beta), when entering recovery mode, or by opting out and subsequently opting back into UMA consent".
[5] Data Source Information Provided Reconciliation - Brown.xlsx

6. UMA user action events record every user interaction with the Chrome browser along with the exact time of each of these interactions. These events include window and tab open, close, and switching events as well as searches and page loads. Some of the user action events pertaining to Incognito usage are listed in GOOG-CABR-00000015 at -22 and -23. These include an [redacted] action which marks the start of an Incognito session and an [redacted] action which marks the end of an Incognito session.[6] Thus, within Google's raw UMA logs, Google keeps track of the exact time when a user has an active Incognito session. This data is not only associated with the user's UMA ID but also with the IP address and the user agent of the user's device.[7]

7. An example UMA user-action sequence is shown below extracted from Plaintiff William Byatt's UMA data produced from the Second Iterative Search of the Special Master process.[8] This user-action sequence indicates the start of an Incognito session along with the timestamp of when the Incognito actions occurred. In addition, the UMA data file contains Mr. Byatt's IP address, and user-agent.[9]

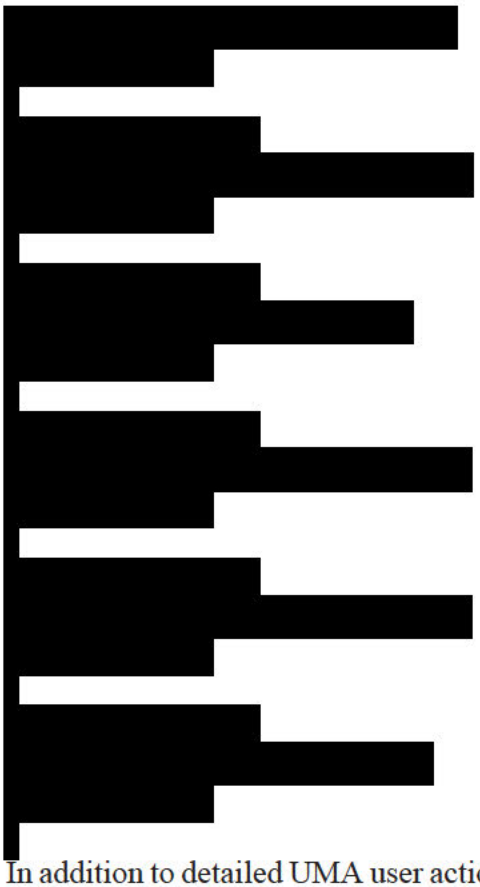
[redacted]

[6] GOOG-CABR-00000015 at -22 and -23 shows many other Incognito related user actions, including several that are specific to the Chrome browser app on mobile devices.

[7] 2021-11-18 Brown Omnibus Sheet.xlsx tab '3.6 [redacted]'. The user's IP address is stored at least in the [redacted] and [redacted] fields. The user agent is stored in the [redacted] field.

[8] The original produced UMA data is contained within [redacted] produced on March 25, 2022 as a part of the Second Iterative Search. The field names have been decoded and included in **Exhibit E** titled [redacted] The full decoded file contains [redacted] lines. The short user-action sequence shown above are from lines 932660 to 932686.

[9] For example, line 67128 of **Exhibit E** shows an IP address of [redacted] and a user agent of [redacted] This IP address and user agent pair appears in a [redacted] – e.g., row 24) produced on the same day. As I describe in Appendix H, this log record, with Biscotti ID [redacted] is associated with Mr. Byatt's Incognito browsing activities. The same IP address and user agent pair also appears in Mr. Byatt's Google account information associated with his GAIA ID and e-mail addresses as shown in Exhibit D in file "wmjbyatt.SubscriberInfo.html".



8. In addition to detailed UMA user actions, Google also obtains aggregated information from each user's Chrome browser in the form of Histograms. Some of the Histogram events pertaining to Incognito usage are listed in GOOG-CABR-00000015. In particular, Google has a webpage load-counting metric called [redacted], which is considered by Google engineers as the "most precise measure of incognito usage."[10] In mid-2020, the name of this metric was changed to [redacted].[11] This metric is used to separately report the number of webpage loads in Incognito and non-Incognito modes during each UMA upload

[10] GOOG-BRWN-00455104 "I think the most precise measure of incognito usage is based on [redacted] UMA."

[11] GOOG-BRWN-00183909. Google engineer explains at https://source.chromium.org/chromium/chromium/src/+/main:tools/metrics/histograms/metadata/navigation/histograms.xml (Last accessed on April 11, 2022) that the [redacted] metric is "The browser profile type for each main-frame navigation, recorded after navigation completion, including NTP [New Tab Page]." [redacted] is specified at https://source.chromium.org/chromium/chromium/src/+/main:components/profile_metrics/browser_profile_type.h?q=BrowserProfileType (Last accessed on April 11, 2022). GOOG-CABR-03717247 at -253 explains histogram buckets.

interval. [12] For example, Mr. Byatt's UMA data contains the histogram event [redacted].[13]

[redacted]

9. Aside from the raw UMA log [redacted], Google also keeps aggregated UMA data in other logs such as [redacted], [redacted], and [redacted], which include aggregated data processed from the raw log.[14] The [redacted] log retains approximately [redacted] of data while the [redacted] logs are retained [redacted].[15] These logs are used for generating aggregate statistics across all users that upload UMA data, such as the number of page loads in Incognito mode and the number of Incognito users. Google uses an UMA Dashboard to extract these aggregate Incognito usage statistics (*see e.g.*, GOOG-CABR-05876937).

[12] GOOG-CABR-04486714 showing the raw page load counts in Chrome regular and Incognito modes based on [redacted]

[13] Taken from lines 1034962 to 1024973 of **Exhibit E**. Each UMA upload contains [redacted] and [redacted] which marks the time duration of histogram counts (e.g., lines 1090087 to 1090096 of **Exhibit E**).

[14] GOOG-BRWN-00032906 at -906

[15] GOOG-BRWN-00032906 at -906 and GOOG-CABR-04801490: "The data shown on UMA dashboards is powered by processed tables - usually either the Timeline table (which will continue to keep indefinitely since it does not have client ids) or the full dynamic table [redacted] which will continue to have [redacted] retention."

10. UMA is considered by Google engineers to be the "ground-truth" for Chrome usage and Incognito statistics.[16] "UMA is the only reliable source for incognito usage statistics at Google…If someone asked me how to find out about incognito usage statistics, I would tell them '[Y]ou should use UMA'" (Sadowski Tr. 28:18-23:6-8). Google reports aggregated 7-day and 28-day statistics based on UMA, including the number of page loads, sessions, and unique users in Incognito and non-Incognito modes across desktop, Android, iOS, and other device platforms.[17]

11. Google's raw UMA log, [redacted], contains data that can readily be used to identify members of the Chrome class. Those class members selected to upload their UMA data can furnish their UMA ID [redacted] and Google account email for verification. UMA data from the previous [redacted] is stored in the raw UMA log if the user's device was selected for UMA upload. I understand from Counsel that Google represented to the Special Master that the types of data stored within the [redacted] are the same as more recent data.

12. Once a Google account holder's UMA data has been located in the raw UMA log, the following information can be obtained:

- User's IP address and user-agent. The geographical location of the user (e.g., United States) can be obtained from the IP address.
- The number of Incognito page loads for each upload period as well as the time duration within which the pageloads were measured.
- The start and end times of every Incognito session along with user browser interaction events within each Incognito session.

13. These metrics identify a Google account holder as an Incognito user in the United States within the previous [redacted] and her Incognito page load counts during those sessions. Combined

[16] GOOG-CABR-03849022 at -022 Google engineer Bert Leung says: "we do have more updated stats from UMA…I believe we can use it as the "ground truth" for analysis." and GOOG-BRWN-00438510 at -510: "in the area of measuring incognito usage I'd strongly suggest believing the UMA data".
[17] See e.g., GOOG-BRWN-00164509, GOOG-CABR-04483323, GOOG-CABR-04648434, GOOG-BRWN-00164509

with the fact that users rarely sign in to their Google account in Incognito mode[18] and the fact that Google tracking beacons are so prevalent on the Internet, the probability of an Incognito user identified in this way being a member of the Chrome class is close to certain.

14. The time-stamped UMA data can be used in conjunction with private browsing information stored in Google ads and Analytics logs to confirm that the user is a member of the class. The ads and Analytics logs can be used to identify specific non-Google websites the user visited during each Incognito session while not signed in to Google. This can be done as follows:

- Identify the relevant log sources through either: (1) IP address and user agent; (2) User's non-Google website sign-in identifiers (such as PPID, Analytics User ID, or Google Ads User ID); (3) Biscotti cookies; (4) Analytics CIDs; or (5) any combination or all of the above.
- For each relevant log source, identify timestamped events stored in the log that took place within each Incognito session duration as identified by UMA. If any such event is found in signed-out logs, then the user has visited a non-Google website while not signed in to Google. Some of these logs also contain explicit Incognito signals tied to individual events as I will discuss in the next subsection.

APPENDIX G.2 INCOGNITO SIGNALS IN LOGS

15. It is my opinion that Google can also use its records of advertising data to identify members of Class I. As with the UMA method described above, this method is being impaired by Google's ongoing deletion of private browsing information. I understand from Counsel that Google has

[18] GOOG-CABR-04959606 at 609: "Incognito: Users in incognito mode usually don't sign-in with gaia" and GOOG-BRWN-00035610 at -616: "Even in Incognito mode, some users do choose to sign in and search. Signed in searches do represent a small fraction of all Incognito searches, however."

represented that it did not change the routine deletion period for any log or data source that contains private browsing information.

16. Google has been tracking Incognito usage not only through UMA but also through an Incognito signal called x-client-data that is only accessible by Google.[19] This x-client-data field is only sent in non-Incognito mode from Chrome browsers to Google and is not sent in Incognito mode.[20] Thus the x-client-data field can be used to detect Incognito mode, and indeed Google has been using the x-client-data field for Incognito detection since at least 2014.[21] Several Google ads event logs store x-client-data (or contain an empty field indicating its absence) along with event timestamp, URL visited, IP address, user agent, and a plethora of other information.[22]

---

[19] GOOG-CABR-00890415 at -417: "This is the X-Client-Data header which gets sent to google as well as some of these domains: {". android.com ", ". doubleclick.com ", ". doubleclick net ", ". ggpht.com ", ". googleadservices.com ", ". googleapis.com ", ". googlesyndication.com ", ". googleusercontent.com ", ". googlevideo.com ", ". gstatic.com ", ". ytimg.com ""} and GOOG-BRWN-00035610 at -610: "Finch is the Chrome experiments framework. In order to facilitates understanding the impact of Chrome experiments on user behavior on Google properties, Finch experiments can report their state to certain Google top level domains. This is accomplished by sending back an additional HTTP header – the X-Client-Data header (which was called X-Chrome-Variations before M33). This header contains a serialized, base64 encoded ChromeVariation proto"

[20] Absent Google code, which they have refused to provide, I have examined publicly available code for Chromium browsers (upon which the Chrome browser is built) and found that the browser implements logic to only transmit the x-client-data header for non-Incognito profiles in the browser. See https://chromium.googlesource.com/chromium/src.git/+/62.0.3178.1/components/variations/net/variations_http_headers.cc?autodive=0%2F%2F%2F (Last accessed on April 11, 2022).: "Note the criteria for attaching client experiment headers: 1. We only transmit to Google owned domains which can evaluate experiments…2. Only transmit for non-Incognito profiles." This is validated by Google documentations (e.g., GOOG-BRWN-00035610 at -610 "These headers are ONLY transmitted when the request is not coming from an Incognito windows").

[21] *See e.g.*, GOOG-BRWN-00035610 at -610: "we can determine the fraction of Chrome searches from Incognito windows by looking at searches that have HTTP headers logged but do not have the X-Client-Data", GOOG-BRWN-00536949 "Incognito traffic from Chrome on Android can be detected from the absence of an X-Client-Data header in search requests" and GOOG-BRWN-00204684 on "How to Detect Incognito Mode" demonstrating Incognito detection for both search and display ads traffic.

[22] Through the Special Master process, three logs containing x-client-data have been identified, including [redacted], [redacted], and [redacted]. Two of these logs have been used by Bert Leung in his Incognito detection study. *See,* GOOG-CABR-05280756 at -756: "Right now log analysis of Chrome Incognito is only possible for logs with both UserAgent and HttpHeaders logged. While UserAgent is generally logged, HttpHeaders are not (It's only logged in a few log sources.). In go/[redacted]-monitor, I am using [redacted] and [redacted]." The Plaintiffs have repeatedly requested Google to disclose other logs containing x-client-data. Google has not responded to those requests. On March 17, 2022, Google disclosed for the first time three additional logs that Google has reviewed, analyzed, or searched as part of Google's efforts to conduct a "log analysis of Chrome Incognito" in and around mid-2020: [redacted], and [redacted] (SUPPLEMENTAL RESPONSE TO INTERROGATORY NO. 34). Google has not provided schema for these logs and it is unclear whether these logs contains x-client-data header information.

17. In 2014, a Google internal document shows Google using the x-client-data header to "determine the fraction of Chrome searches from incognito window by looking at searches that have HTTP headers logged but do not have the X-Client-Data" (GOOG-BRWN-00035610 at -610). This effort included an identification of user-agent from browsers that "are not capable of sending the X-Client-Data header" (GOOG-BRWN-00035610 at Table 1).

18. In 2020, just before this lawsuit was filed, Google engineers were engaged in detecting Incognito traffic from traffic Google intercepted, collected, and stored in ads event logs. By looking at the x-client-data field in various logs, Google engineers were able to determine Incognito usage with a high degree of accuracy. This effort culminated in a Dashboard[23] that Google engineering teams used to monitor Incognito traffic starting at least in the middle of 2020.

19. Subsequently, Google engineers made improvements on the Incognito detection logic and developed a more sophisticated [redacted] Monitoring Dashboard. This follow-on work included logging a "maybe_chrome_incognito" field into [redacted] event logs, including four signed-out Display ad logs.[24] Google engineers initially referred to this field as the "is_incognito" field and changed the name to "maybe_chrome_incognito" in July 2020 after this lawsuit began.[25]

20. Elsewhere in the Google infrastructure, x-client-data-based Incognito fields are referred to as "is_chrome_incognito" and "is_chrome_non_incognito." [26] These two fields have been logged in [redacted] logs since 2017.[27] In addition to relying on the absence of the x-client-data header, the "is_chrome_incognito" and "is_chrome_non_incognito" fields may also rely on the absence of the X-Geo header (GOOG-BRWN-00536949 discussing [redacted] logs and noting that "Chrome

[23] GOOG-CABR-03662990 at -992
[24] 2022-03-01 Brown v. Google - Google Letter re [redacted] Monitoring Dashboard.pdf
[25] GOOG-CABR-00547295 at -296
[26] See e.g., Final Sadowski Reference Sheet listing several [redacted] logs containing is_chrome_incognito and is_chrome_non_incognito fields
[27] *Id.*

on Android attaches location to Omnibox search queries via a header (X-Geo header) except in Incognito mode."; Sadowski Tr. 71:8-23).

21. As Google engineers explain in a document titled "How to Detect Incognito Mode?", "The starting point will be the (absence of) X-Client-Data header. Chrome uses Finch (go/finch) framework for A/B experiments. It sends experiment IDs via the special HTTP header X-Client-Data to Google servers (including doubleclick.net). While Chrome doesn't send this header in Incognito Mode, there are other cases where Chrome also doesn't send this header and that therefore need to be excluded as well - the majority of which are Android Webviews and other Chromium-based browsers."[28]

22. While there are limited cases where Chrome does not send x-client-data for non-Incognito traffic, the Google engineers were able to apply filters to eliminate many of the false positives (i.e., a determination of Incognito traffic based on the absence of x-client data when in fact the traffic was not generated in Incognito mode).

23. Throughout the month of May 2020, Google engineer Bert Leung and others were able to progressively reduce the Incognito detection error rate for their Incognito detection analysis as they sought to achieve a statistic consistent with UMA pageload data.[29] Mr. Leung's colleague Chris Liao remarked that "I think these are very good results indeed."[30] Their presentation in June 2020 stated: "Incognito detection accuracy improved materially over the past two weeks". When compared to UMA Incognito statistics, which Google engineers considered to be the ground truth as discussed earlier, Incognito traffic from log-based analysis was [redacted] compared to UMA

[28] GOOG-CABR-04485508 at -508

[29] See e.g., GOOG-CABR-00546721 at -721: "Update: I have driven the number down to [redacted] using a strict regex based filter…I believe this method should be pretty much the limit of current methodology and should be good enough for short-term monitoring." (May 28. 2020). GOOG-CABR-03663606 at '609: "If we use a heuristic to de-dup spam clicks by query_id, …overall incognito request % drops down to [redacted] based on [redacted] of data, further closer to Chris' number."

[30] GOOG-CALH-00376260 at -510

Incognito data at [redacted].[31] Google engineers then improved on this result by eliminating webview traffic[32], which was the main source of error.[33] Engineer Mandy Liu reported to her manager, Bert Leung, "Good news: Chrome incognito rate is [redacted]!" (GOOG-BRWN-00846508 at -508). Ms. Liu testified that this was "good news" because this result was "closer to the number Bert [Leung] told me" (i.e., the UMA [redacted] figure), and she explained that Google achieved this result by successfully excluding mobile app traffic from Chrome traffic (Liu Tr. 39:14-16, 40:15-17). Thus, Google engineers achieved extremely high accuracy with their Incognito detection logic, which can also be used to identify class members.

24. The "maybe_chrome_incognito" field is now being used by a Google dashboard (Factual Stipulation re: maybe_chrome_incognito). With the "maybe_chrome_incognito" field already good enough for Google's internal use, any residual error rate on an event-by-event basis is merely the worst one can do based on the specific scenario Google analyzed. For the purpose of class identification, x-client-data-based Incognito detection can achieve even higher accuracy for at least the following reasons:

1. The percentage of Incognito traffic Google engineers reported counts every traffic event independently of other events from the same browsing session. For class identification purposes, we need to look at the total volume of data belonging to the same user stored across several different logs, instead of individual traffic events. Based on Google's storage architecture, all signed-out data from a particular user in a particular Incognito session is keyed to the same Biscotti ID. As long as there exists a single event entry containing x-client-data in any of the x-client-data containing logs, we can be sure that the associated Biscotti ID was generated in Chrome non-Incognito mode. Hence, it does not matter if some of the entries do not contain x-client-data or even if a subset of the x-

[31] GOOG-CABR-04795991 at -996

[32] Webview are embedded browser instances in mobile apps and are not a part of this case.

[33] GOOG-BRWN-00846942 at -942: "It was discovered when building up the [redacted] dashboard (go/bi-dash) that [redacted] V1 traffic rate is exceptionally high. After looking into some captures, it appears that the majority of the traffic is not identified as WebView, but has ; wv in their user agents (example). [redacted] V1 detection for Display Ads traffic relies on WebView detection in [redacted], as WebView traffic should be excluded. But the [redacted] signal we are using to determine [redacted] V1 only reflects requests from web WebView (i.e. WebView as a page containing ads), leaving the ones from app WebView (i.e. WebView as an ad slot)."

client-data containing logs do not contain x-client-data. As long as there exists a single event entry with x-client-data, the associated Biscotti ID was generated in non-Incognito mode.

2. Google's Incognito traffic reports are for world-wide traffic, not limited to traffic from the United States. Google has not provided an error rate study specifically for traffic from the United States, but according to Bert Leung, some of the false positives are due to traffic from other countries (e.g., Japan) (GOOG-CABR-03662990 at -991).
3. In July 2020, Chrome launched User Agent Client Hints (UACH).[34] This is another HTTP Header field that explicitly states that "Google Chrome" is the browser used, and is distinguished from other browsers. Thus, at least since July 2020, Chrome traffic can be further distinguished from other Chromium browser traffic using UACH, thereby eliminating much of the false positives from other non-Chrome browsers.
4. For users with UMA data uploaded, event data logs can be matched with UMA data logs using IP address, user agent, and timestamp to further demonstrate Incognito usage at a particular time by a particular user.

25. In Appendix H.11 and Appendix I, I discuss using x-client-data in ads logs to identify Incognito traffic and show that for data from the Second Iterative Search, Incognito detection accuracy is 100%. Furthermore, Google's "maybe_chrome_incognito" field is also 100% accurate for that set of data.

26. Through the Special Master process, I have observed x-client-data in Plaintiffs' data. For example, Biscotti ID [redacted] is associated with a Biscotti cookie collected from Mr. Byatt's Chrome regular mode session, and Biscotti ID [redacted] is associated with a Biscotti cookie collected from Mr. Byatt's Chrome Incognito mode session. Comparing the display ads interaction data that Google collected and stored for these two Biscotti IDs,[35] it is clear that x-client-data sent from the browser is logged in Chrome regular mode and not present in

[34] GOOG-CABR-00398813 at -814 UA-CH is scheduled to be release to stable on July 14th with Chrome M84"
[35] E.g., in the "2022-03-14 Brown v. Google - [redacted] – AEO" production, within [redacted] and [redacted]

Chrome Incognito mode.[36] Both the Incognito and non-Incognito mode logs store Mr. Byatt's IP address, user-agent, and time stamp for ads events.

27. Google engineer Dr. Glenn Berntson identified four scenarios he described as "false positives" in which the x-client-data header would not be sent within a non-Incognito session. These four scenarios range from merely theoretical to exceedingly rare.

28. First, Dr. Berntson testified that the x-client data header would not be sent for a "new browser instance" (Berntson June 16 Tr. 379:6-7). But he clarified that this would only apply until "the call that Chrome makes to retrieve the variation IDs is successful" (*Id.* 379:10-12). And this process would be completed almost instantaneously. At worst, "if it's a slow network connection it can [] taker a minute or two" (*Id.* 380:14-16).

29. The second "false positive" scenario described by Dr. Berntson is when a browser has been inactive for 30 days or more—the x-client data header "is considered to be stale and just purged and no x-client data header is passed." *Id.* 375:8-12. But Dr. Berntson conceded that, like with his first scenario, this scenario would, at worst, apply only for a minute or two (based on a slow connection). *Id.* 381:22-23.

30. The third example he gave is "if the variation IDs that are carried in the x-client data header, if too many are returned to Chrome to prevent the requests coming from Chrome from being too large, they just delete them all and so you'd see no x-client data header." *Id.* 375:13-19. But when directly asked, "how often" this occurs, he admitted that "I have not seen this flagged as a problem." *Id.* 385:4-5.

---

[36] In logs associated with Incognito data, two different x-client-data fields sometimes appear together within the same event entry: "[redacted]" and "[redacted]" (See e.g., row 115 of [redacted]. Within such Incognito log, the [redacted] field is always empty, which accurately reflects that the browsing activities are from Incognito mode where x-client-data header is not sent. In contrast, logs associated with Chrome Regular mode data contain x-client-data sent from the browser (See e.g., row 16 of [redacted] showing "[redacted]").

31. Dr. Berntson's fourth example was "the presence of a firewall," which he said "can also prevent Chrome . . . from getting the variation IDs that are used to populate the X-Client Data header." Id. 375:20-376:11. This, he explained, was due to a particular Google server endpoint being blocked by the firewall. Dr. Berntson did not identify the particular Google server endpoint and whether this server endpoint has been identified by any firewall to be a security risk. The reality is that Google engineers were able to detect Incognito traffic with an extremely high degree of accuracy. Any hypothetical error from a firewall blocking particular Google servers is unlikely to materially impact Incognito identification.

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

APPENDIX H

Appendix H Data Analysis

Content

1. Within the main body of my report, I have presented a few examples of Google's logs and joinability risks using data produced from the Special Master process. In this Appendix, I provide additional supporting information for those examples. In addition, I demonstrate in Section H.11 that using Google's Incognito detection signals in logs achieves 100% accuracy for data produced in the Second Iterative Search.

Appendix H.1 Plaintiff Chasom Brown's Incognito User Profile in [redacted]

2. From the "2022-03-14 Brown v. Google - DBL [redacted] – AEO" production, [redacted] contains Plaintiff Chasom Brown's Incognito user profile information associated with Biscotti ID [redacted]. Mr. Brown's private browsing information associated with this Biscotti ID is used to build his profile stored in DBL [redacted].

3. The Biscotti ID value [redacted] is a hexadecimal value converted from the decimal Biscotti ID [redacted]. This decimal to hexadecimal value conversion is shown below.[1]

[redacted]

4. Google has represented that this Biscotti ID corresponds to four Biscotti cookies.[2] While the cookie values have changed over time, the embedded uid (Biscotti ID) has not.

Item	Biscotti Cookie	uid	Creation Timestamp
1	[redacted]	[redacted]	[redacted]
2	[redacted]	[redacted]	[redacted]
90	[redacted]	[redacted]	[redacted]
98	[redacted]	[redacted]	[redacted]

5. By way of example, Biscotti cookie (IDE) corresponding to item 2 in the table [redacted] was extracted from Mr. Brown's computer in Incognito mode. In one instance, this Biscotti cookie was sent to Google's doubleclick.net server on [redacted], while visiting a page on [redacted] in Incognito mode and not signed into his

[1] https://www.rapidtables.com/convert/number/decimal-to-hex html (Last accessed on April 11, 2022).
[2] 2022-03-02 Brown v. Google - Decode IDE.pdf

Google account ([redacted]). Note that in Chrome Incognito mode, X-Client-Data was not sent to Google.

Appendix H.2 Plaintiff William Byatt's Incognito User Profile in

6. From the "2022-03-14 Brown v. Google - DBL – AEO" production, contains Plaintiff William Byatt's Incognito user profile information associated with Biscotti ID . Mr. Byatt's private browsing information associated with this Biscotti ID are used to build his profile stored in DBL .

7. The Biscotti ID value is a hexadecimal value converted from Biscotti ID . This decimal to hexadecimal value conversion is shown below.[3]

[3] https://www.rapidtables.com/convert/number/decimal-to-hex html (Last accessed on April 11, 2022).

[redacted]

8. Google has represented that this Biscotti ID corresponds to eight Biscotti cookies.[4] While the cookie values have changed over time, the embedded uid (Biscotti ID) has not.

Item	Biscotti Cookie	uid	Creation Timestamp
9	[redacted]	[redacted]	[redacted]
10		[redacted]	[redacted]
12		[redacted]	[redacted]
13		[redacted]	[redacted]
14		[redacted]	[redacted]
15		[redacted]	[redacted]
91		[redacted]	[redacted]
96		[redacted]	[redacted]

9. By way of example, Biscotti cookie (IDE) corresponding to item 9 in the table [redacted] was extracted from Mr. Byatt's computer in Incognito mode. In one instance, this Biscotti cookie was sent to Google's doubleclick.net server on [redacted], while visiting

[4] 2022-03-02 Brown v. Google - Decode IDE.pdf

[redacted] in Incognito mode and not signed into his Google account (No GAIA DSID cookie).

Note that in Chrome Incognito mode, X-Client-Data was not sent to Google.

Appendix H.3 Plaintiff Jeremy Davis's Incognito User Profile in [redacted]

10. From the "2022-03-14 Brown v. Google - DBL [redacted] – AEO" production, [redacted] contains Plaintiff Jeremy Davis's Incognito user profile information associated with Biscotti ID [redacted]. Mr. Davis's private browsing information associated with this Biscotti ID are used to build his profile stored in DBL [redacted].

11. The Biscotti ID value [redacted] is a hexadecimal value converted from Biscotti ID [redacted]. This decimal to hexadecimal value conversion is shown below.[5]

[redacted]

12. Google has represented that this Biscotti ID corresponds to five Biscotti cookies.[6] While the cookie values have changed over time, the embedded uid (Biscotti ID) has not.

Item	Biscotti Cookie	uid	Creation Timestamp
33	[redacted]	[redacted]	[redacted]
34		[redacted]	[redacted]
35		[redacted]	[redacted]

[5] https://www.rapidtables.com/convert/number/decimal-to-hex html (Last accessed on April 11, 2022).
[6] 2022-03-02 Brown v. Google - Decode IDE.pdf

36	[redacted]	[redacted]	[redacted]
37		[redacted]	[redacted]

13. By way of example, Biscotti cookie (IDE) corresponding to item 35 in the table [redacted] was extracted from Mr. Davis's computer in Incognito mode. In one instance, this Biscotti cookie was sent to Google's doubleclick.net server on [redacted], while visiting a specific page on [redacted] in Incognito mode and not signed into his Google account ([redacted]). Note that in Chrome Incognito mode, X-Client-Data was not sent to Google.

[redacted]

[redacted]

Appendix H.4 Plaintiff Jeremy Davis's Regular Mode User Profile in [redacted]

14. From the "2022-03-14 Brown v. Google - DBL [redacted] – AEO" production, [redacted] contains Plaintiff Jeremy Davis's Chrome regular mode user profile information associated with Biscotti ID [redacted]. Mr. Davis's Chrome regular mode browsing information associated with this Biscotti ID are used to build his profile stored in DBL [redacted].

15. The Biscotti ID value [redacted] is a hexadecimal value converted from Biscotti ID [redacted]. This decimal to hexadecimal value conversion is shown below.[7]

[redacted]

16. Google has represented that this Biscotti ID corresponds to five Biscotti cookies.[8] While the cookie values have changed over time, the embedded uid (Biscotti ID) has not.

Item	Biscotti Cookie	uid	Creation Timestamp
38	[redacted]	[redacted]	[redacted]
39	[redacted]	[redacted]	[redacted]

[7] https://www.rapidtables.com/convert/number/decimal-to-hex html (Last accessed on April 11, 2022).
[8] 2022-03-02 Brown v. Google - Decode IDE.pdf

40	[redacted]	[redacted]	[redacted]
41		[redacted]	[redacted]
42		[redacted]	[redacted]

17. By way of example, Biscotti cookie (IDE) corresponding to item 38 in the table [redacted] was extracted from Mr. Davis's computer in regular mode. In one instance, this Biscotti cookie was sent to Google's doubleclick.net server on [redacted], while visiting [redacted] in Chrome regular mode and not signed into his Google account [redacted]). Note that in Chrome regular mode, X-Client-Data was sent to Google.

[redacted]

Appendix H.5 Plaintiff Monique Trujillo's Regular Mode User Profile in [redacted]

18. From the "2022-03-14 Brown v. Google - DBL [redacted] – AEO" production, [redacted] contains Plaintiff Monique Trujillo's Chrome regular mode user profile information associated with Biscotti ID [redacted] Ms. Trujillo's Chrome regular mode browsing information associated with this Biscotti ID are used to build her profile stored in DBL [redacted].

19. The Biscotti ID value [redacted] is a hexadecimal value converted from Biscotti ID [redacted]. This decimal to hexadecimal value conversion is shown below.[9]

[9] https://www.rapidtables.com/convert/number/decimal-to-hex html (Last accessed on April 11, 2022).

20. Google has represented that this Biscotti ID corresponds to six Biscotti cookies.[10] While the cookie values have changed over time, the embedded uid (Biscotti ID) has not.

Item	Biscotti Cookie	uid	Creation Timestamp
79			
80			
81			
82			
85			
99			

21. By way of example, Biscotti cookie (IDE) corresponding to item 79 in the table

was extracted from Ms. Trujillo's computer in regular mode. In one instance, this Biscotti cookie was sent to Google's doubleclick.net server on , while visiting a page on in Chrome regular mode and not

[10] 2022-03-02 Brown v. Google - Decode IDE.pdf

signed into her Google account ([redacted]). Note that in Chrome regular mode, X-Client-Data was sent to Google.

[redacted]

Appendix H.6 Plaintiff Monique Trujillo's User Profile in [redacted] Associated with PPID-Mapped-Biscotti

22. From the "2022-03-14 Brown v. Google - DBL [redacted] – AEO" production, [redacted] contains Plaintiff Monique Trujillo's user profile information associated with PPID-mapped-Biscotti ID [redacted]. The PPID-mapped-Biscotti value is associated with both Incognito and Chrome regular mode browsing data.

23. The Biscotti ID value [redacted] is a hexadecimal value converted from Biscotti ID [redacted]. This decimal to hexadecimal value conversion is shown below.[11]

[redacted]

[11] https://www.rapidtables.com/convert/number/decimal-to-hex.html (Last accessed on April 11, 2022).

24. Google has represented that this PPID-mapped-Biscotti ID corresponds to PPID: [redacted] from [redacted].[12] This PPID was sent to Google when visiting [redacted] in both Incognito and Regular mode and signed into Mr. Trujillo's [redacted] account.

25. By way of example, PPID: [redacted] was extracted from Ms. Trujillo's computer from both regular and Incognito modes. In one instance, this PPID was sent to Google's doubleclick.net server on [redacted], while visiting [redacted] in Chrome regular mode, signed into her [redacted] account, and not signed into her Google account ([redacted]). Note that in Chrome regular mode, X-Client-Data was sent to Google.

[redacted]

[12] 2022-03-04 Brown Second Iterative Searches – PPID.xlsx

26. In another instance, the same PPID: [redacted] was sent to Google's doubleclick.net server on [redacted], while visiting [redacted] in Incognito mode, signed into her [redacted] account, and not signed into her Google account (No GAIA cookie). Note that in Chrome Incognito mode, X-Client-Data was not sent to Google.

Appendix H.7 Plaintiff William Byatt's Incognito Ads and Analytics Data

27. In Appendix H.2, I have discussed Mr. Byatt's Biscotti ID: [redacted] as generated from Incognito mode. By way of another example, Biscotti cookie [redacted] was extracted from Mr. Byatt's computer in Incognito mode. In one instance, this Biscotti cookie was sent to Google's doubleclick.net server on [redacted], while visiting a page on [redacted] in Incognito mode and not signed into his Google account (No GAIA DSID cookie). Note that in Chrome Incognito mode, X-Client-Data was not sent to Google.

[redacted]

28. In another instance, the same Biscotti cookie was sent to Google along with Google Analytics CID: [redacted] from [redacted]. For example, both identifiers were sent to Google's doubleclick.net server on [redacted], while visiting [redacted] in Incognito mode and not signed into his Google account (No GAIA DSID cookie). Note that in Chrome Incognito mode, X-Client-Data was not sent to Google.

Appendix H.8 Plaintiff Jeremy Davis's Incognito Search and Display Ads Data

29. From the "2022-03-14 Brown v. Google - [redacted] – AEO" production, [redacted] contains Plaintiff Jeremy Davis's Incognito search data based on a search of his Zwieback ID: [redacted] Mr. Davis's Incognito search data associated with this Zwieback ID are stored in [redacted], which is one of the logs Google used for their 2020 log-based Incognito detection study.

30. Google has represented that this Zwieback ID corresponds to eight Zwieback cookies.[13] While the cookie values have changed over time, the embedded uid (Zwieback ID) has not.

Row	Zwieback Cookie	uid
50	[redacted]	
51	[redacted]	[redacted]
56	[redacted]	[redacted]
57	[redacted]	[redacted]
191	[redacted]	[redacted]
192	[redacted]	[redacted]
193	[redacted]	[redacted]

[13] 2022-03-04 Brown v. Google - Second Iterative Searches - Decrypted NIDs.xlsx

194	[redacted]	[redacted]

31. By way of example, Zwieback cookie corresponding to row 50 in the table [redacted] was extracted from Mr. Davis's computer in Incognito mode. In one instance, this Zwieback cookie was sent to Google's [redacted] server on F[redacted], while searching on [redacted] in Incognito mode and not signed into his Google account (No GAIA cookie). Note that in Chrome Incognito mode, X-Client-Data was not sent to Google.

[redacted]

[redacted]

32. From the "2022-03-14 Brown v. Google - [redacted] – AEO" production, [redacted], and [redacted] contain Plaintiff Jeremy Davis's Incognito browsing data based on a search of his Biscotti ID: [redacted] Mr. Davis's private browsing information associated with this Biscotti ID are stored in [redacted], [redacted], [redacted] and [redacted]. I have shown, in Appendix H.3, that Mr. Davis's Biscotti ID [redacted] was extracted from Incognito mode while he was not signed into Google.

Appendix H.9 Consultant Dr. Dai's Incognito Search and Display Ads Data

33. From the First Iterative Search, GOOG-BRWN-00826529 [redacted], GOOG-BRWN-00826530 [redacted], GOOG-BRWN-00826531 [redacted], GOOG-BRWN-00826532 [redacted], GOOG-BRWN-00826534 [redacted], GOOG-BRWN-00826535 [redacted], GOOG-BRWN-00826536 [redacted], GOOG-BRWN-00826537 [redacted] and GOOG-BRWN-00840745 [redacted] contain Dr. Dai's Incognito signed-out data associated with Biscotti ID [redacted].

34. Google has represented that this Biscotti ID corresponds to three Biscotti cookies (GOOG-BRWN-00825709). While the cookie values have changed over time, the embedded uid (Biscotti ID) has not.

Row	Biscotti Cookie	uid
2	[redacted]	[redacted]
5	[redacted]	[redacted]
6	[redacted]	[redacted]

35. By way of example, Biscotti cookie corresponding to row 6 in the table [redacted] was extracted from Dr. Dai's computer in Incognito mode. In one instance, this Biscotti cookie was sent to Google's doubleclick.net server on [redacted], while visiting a page on [redacted] in Incognito mode and not signed into her Google account ([redacted]). Note that in Chrome Incognito mode, X-Client-Data was not sent to Google.

[redacted]

[redacted]

36. From the First Iterative Search, GOOG-BRWN-00826130 ([redacted]) contains Dr. Dai's Incognito signed-out Google search data associated with Zwieback ID [redacted]. Google has represented that this Zwieback ID corresponds to two Zwieback cookies (GOOG-BRWN-00825708). While the cookie values have changed over time, the embedded uid (Zwieback ID) has not.

Row	Zwieback Cookie	uid
372	[redacted]	[redacted]
373	[redacted]	[redacted]

37. By way of example, Zwieback cookie corresponding to row 373 in the table [redacted] was extracted from Dr.

Dai's computer in Incognito mode. In one instance, this Zwieback cookie was sent to Google's [redacted] server on [redacted], while searching on [redacted] in Incognito mode and not signed into her Google account (No GAIA cookie).

Appendix H.10 Plaintiff Jeremy Davis's Regular Mode Search and Display Ads Data

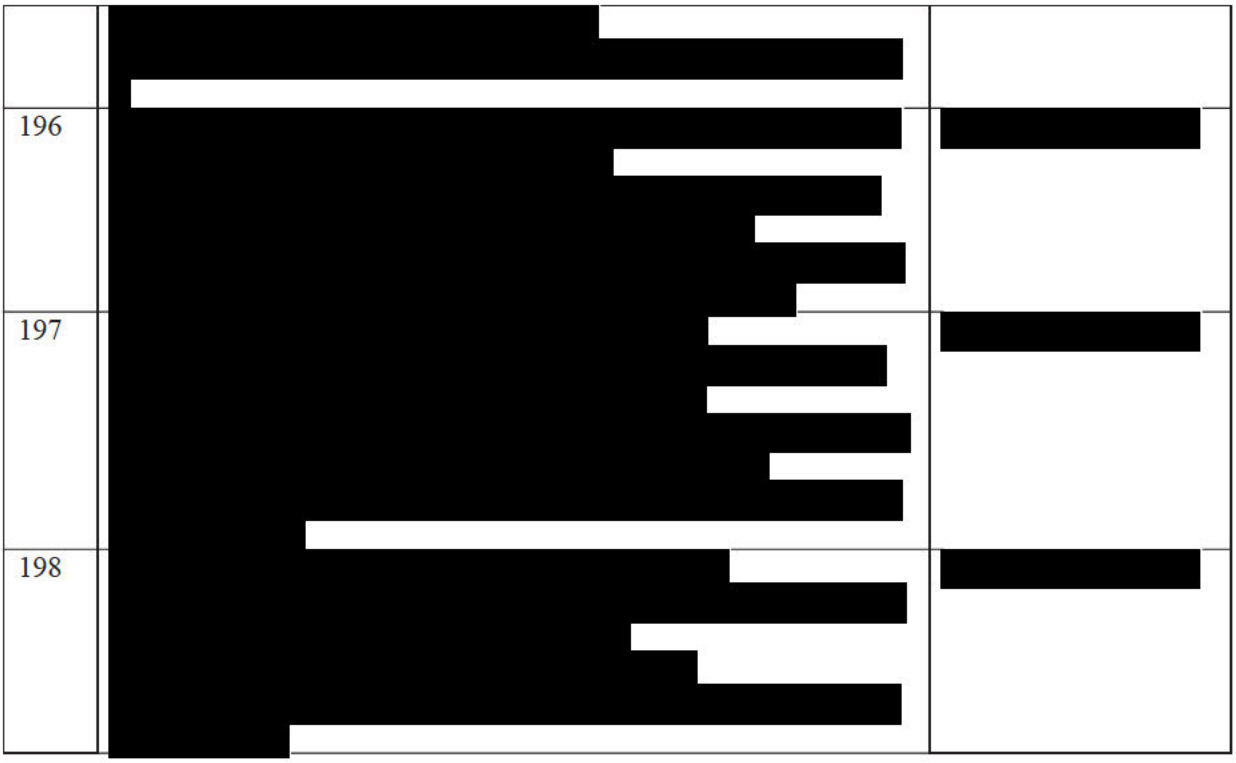
38. As I have shown in Appendix H.4, Biscotti ID [redacted] is associated with Plaintiff Jeremy Davis's Chrome regular mode data.

39. From the "2022-03-14 Brown v. Google - [redacted] – AEO" production, [redacted] contains Plaintiff Jeremy Davis's regular mode search data based on a search of his Zwieback ID: [redacted].

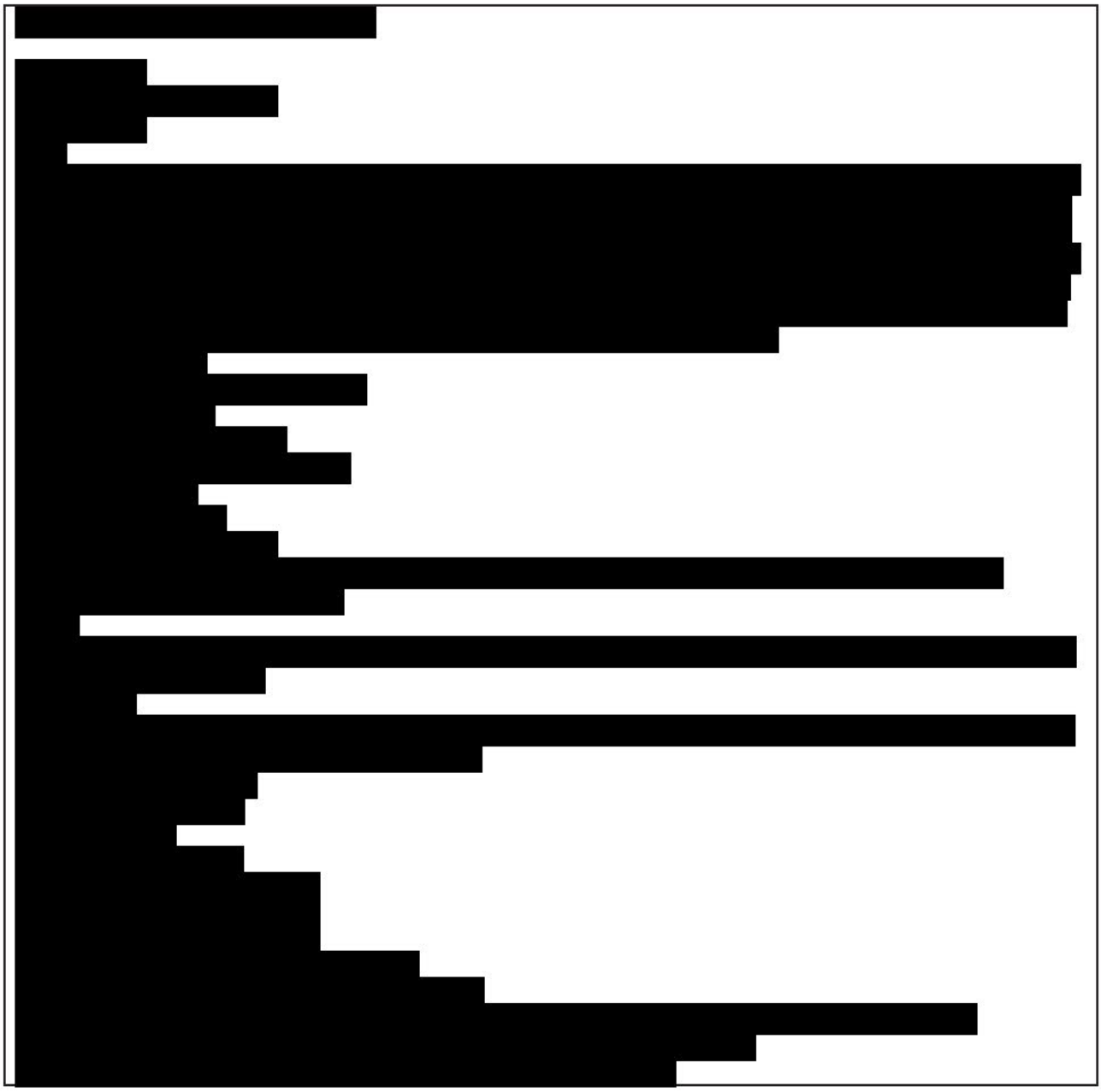
40. Google has represented that this Zwieback ID corresponds to eleven Zwieback cookies.[14] While the cookie values have changed over time, the embedded uid (Zwieback ID) has not.

Row	Zwieback Cookie	uid
52	[redacted]	[redacted]
53	[redacted]	[redacted]
54	[redacted]	[redacted]
55	[redacted]	[redacted]
58	[redacted]	[redacted]
189	[redacted]	[redacted]
190	[redacted]	[redacted]
195	[redacted]	[redacted]

[14] 2022-03-04 Brown v. Google - Second Iterative Searches - Decrypted NIDs.xlsx



41. By way of example, Zwieback cookie corresponding to row 52 in the table [redacted] was extracted from Mr. Davis's computer in regular mode. In one instance, this Zwieback cookie was sent to Google's [redacted] server on [redacted], while searching on [redacted] in regular mode and not signed into his Google account (No GAIA cookie). Note that in Chrome regular mode, X-Client-Data was sent to Google.



Appendix H.11 Incognito Detection for Data Produced in Second Iterative Search

42. During the Second Iterative Search of the Special Master process, Plaintiffs submitted Biscotti cookies values for data search across several Google display ads logs. Google represented that these Biscotti cookies correspond to [redacted] unique Biscotti IDs.[15]

43. On March 4, 2022, during a live demo session, Google engineers produced maybe_chrome_incognito field values for [redacted] display ads logs [redacted] for a small number of Biscotti IDs.[16] maybe_chrome_incognito = true means that a particular event is browsing activity in Incognito mode and maybe_chrome_incognito = false means that a particular event is browsing activity in non-Incognito mode. I have summarized these results in Appendix I, Results tab, columns M to P. For all eight Biscotti IDs with "maybe_chrome_incognito" data, I note that for this sample, the accuracy is 100% (i.e., every event produced from these logs is accurately associated with the Incognito mode or regular mode browsing and that there are no false positive or false negative results).[17] As of April 11, 2022, Google has yet to produce full responsive search results for all of the "maybe_chrome_incognito" logs. I reserve my right to supplement this report should Google produce more data at a later date.

44. I have also analyzed log data for [redacted] ads logs containing x-client-data: [redacted] and [redacted]. Following Google

[15] 2022-03-02 Brown v. Google - Decode IDE.pdf
[16] 2022-03-04 Brown v. Google - Live [redacted] Queries.xlsx
[17] The accuracy of "maybe_chrome_incognito" is in comparison with the actual browsing mode for each Biscotti ID. Each Biscotti cookie is extracted from either regular mode or Incognito mode and the decrypted Biscotti ID is thus associated with either regular mode or Incognito mode as indicated. I have included HTTP extractions taken from browsers on the client side in Appendices H and I. The full HTTP traffic captures are included in **Exhibit F**.

engineers' algorithm for Incognito detection using x-client-data and user agent, I conducted Incognito detection for each Biscotti ID.[18]

45. First, for every event entry in these three logs that contain data, x-client-data information is extracted and shown in Appendix I, "[redacted] Ads XCD Logs" tab along with IP address and user agent. There are three variants of x-client-data in Google logs: (1) lower case x-client-data, (2) X-Client-Data with capitalization, and (3) X-Google-GFE-Original-X-Client-Data as shown in columns C to E of the "[redacted] Ads XCD Logs" tab. A "null" value indicates that the particular variant of x-client-data does not appear in the event entry; an empty cell indicates that the particular variant of x-client-data appears in the event entry but does not contain a value. In particular, X-Client-Data and X-Google-GFE-Original-X-Client-Data sometimes appear in the same event entry. In that case, X-Google-GFE-Original-X-Client-Data stores the true x-client-data value sent from the browser.

46. Next, IP addresses that do not correspond to submitted IP addresses are marked in column H with a value of "0".[19] For the remaining log entries marked in column H with a value of "1", I marked in Column I "Yes" if either the "x-client-data" or "X-Google-GFE-Original-X-Client-Data" contains a value other than "null" or empty; and "No" otherwise. Column J simply contains a "1" if column I is "Yes" and "0" otherwise for ease of computation.

47. A summary of the number of events containing x-client-data is shown in column N for each Biscotti ID. Column N contains the number of event entries that have valid IP address and Column P contains the total number of event entries for each Biscotti ID. In Column Q, I determine

[18] The Incognito detection steps outlined in GOOG-BRWN-00204684 at -685 are: Absence of x-client-data header, Chrome browser, not Webview and not iOS.

[19] It appears that some log events contain IP addresses that belong to Google or other data servers. One of the IP addresses [redacted] appears in an internal document GOOG-CABR-04204890 at -901 that discusses test queries.

the Biscotti ID to be from Incognito mode if Column N contains "0" (i.e., no x-client-data) and from Regular mode if Column N contains a positive integer (i.e., has x-client-data). Two of the Biscottis are from iPhone devices. In keeping with Google's Incognito detection to exclude iOS devices, I have marked these two Biscottis as "iPhone".

48. Incognito detection results based on x-client-data are shown in the "Results" tab in Column I. Comparing this detection result with the actual browsing mode to be Incognito or regular mode in columns D and E, one can see that the detection accuracy is 100% across this data subset.

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

APPENDIX I

Produced Electronically

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

EXHIBIT A

Curriculum Vitae

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

EXHIBITS B1, B2, AND B3

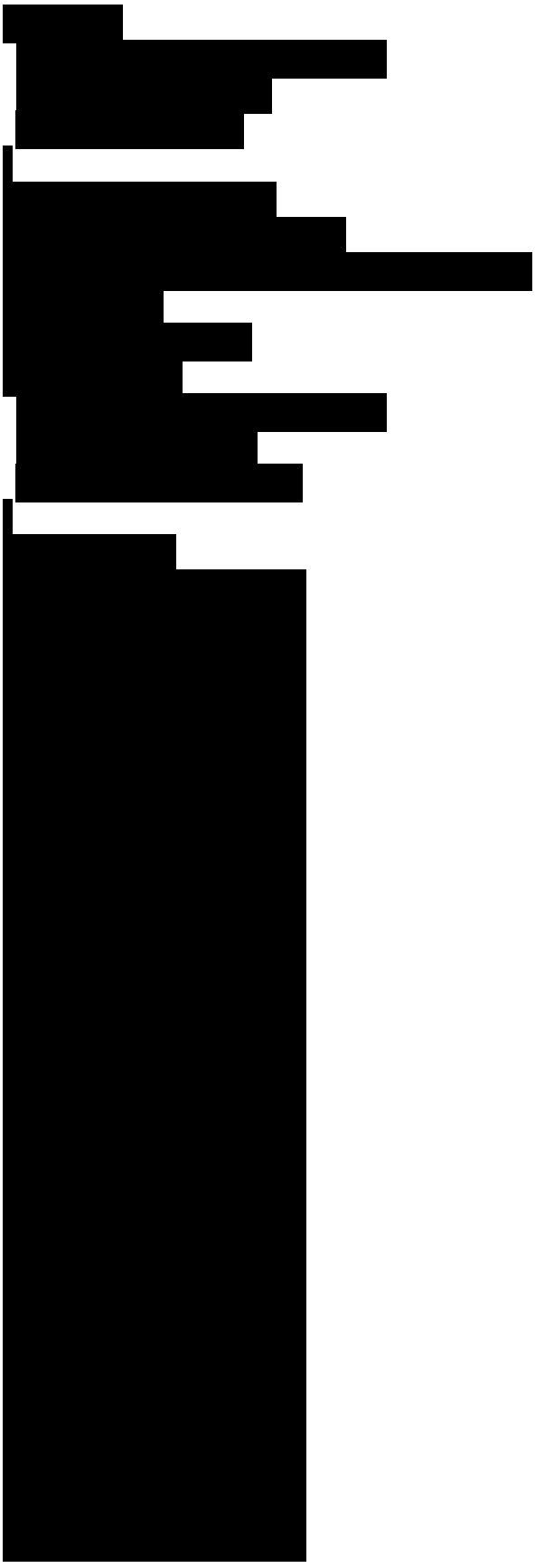
Produced Electronically

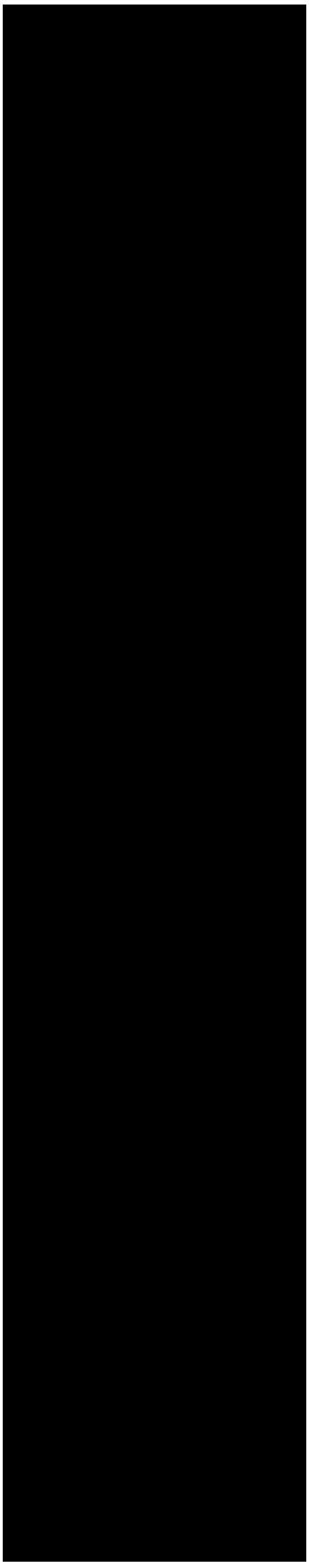
EXPERT REPORT OF JONATHAN E. HOCHMAN

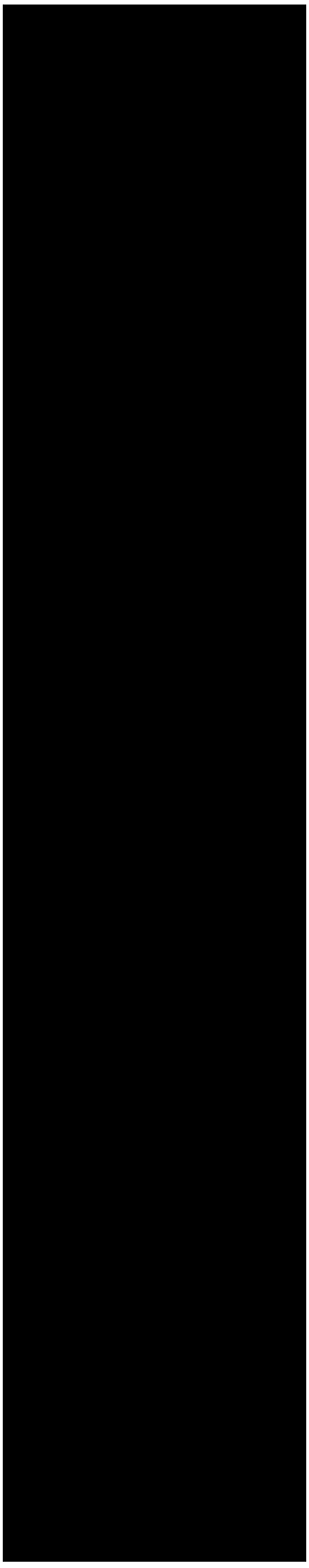
APRIL 15, 2022

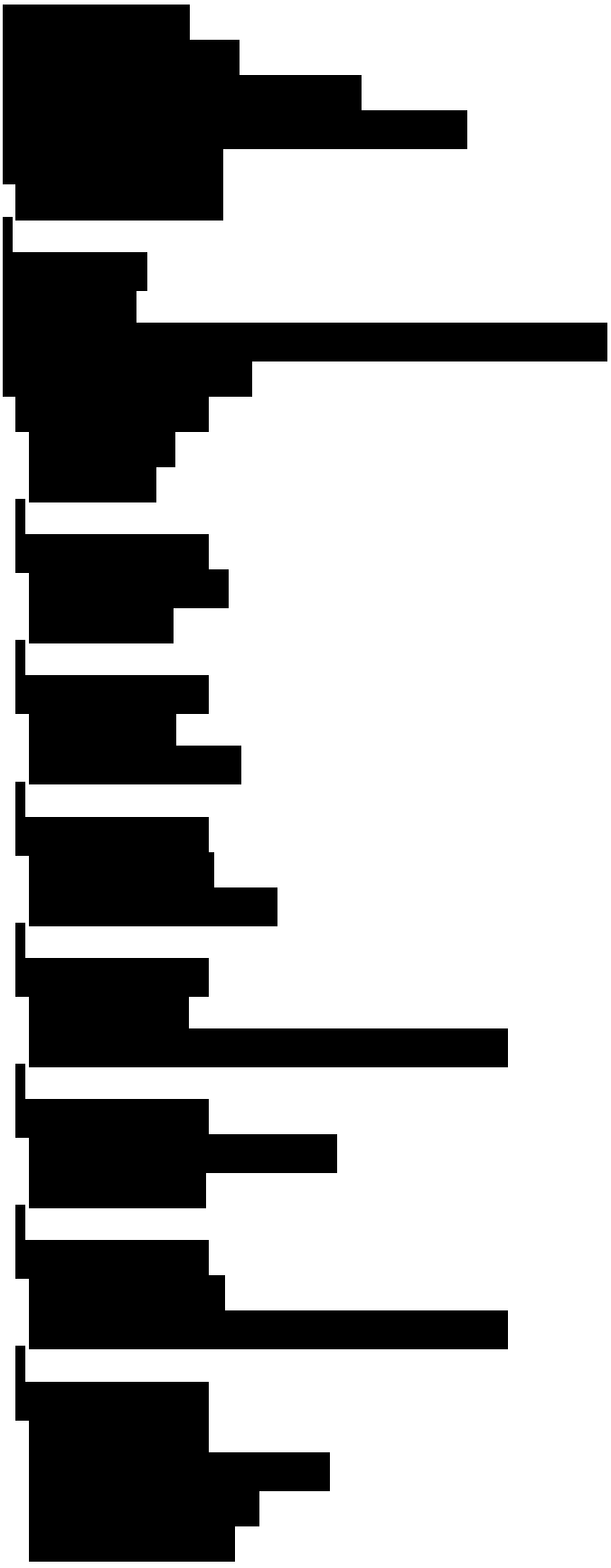
EXHIBIT C

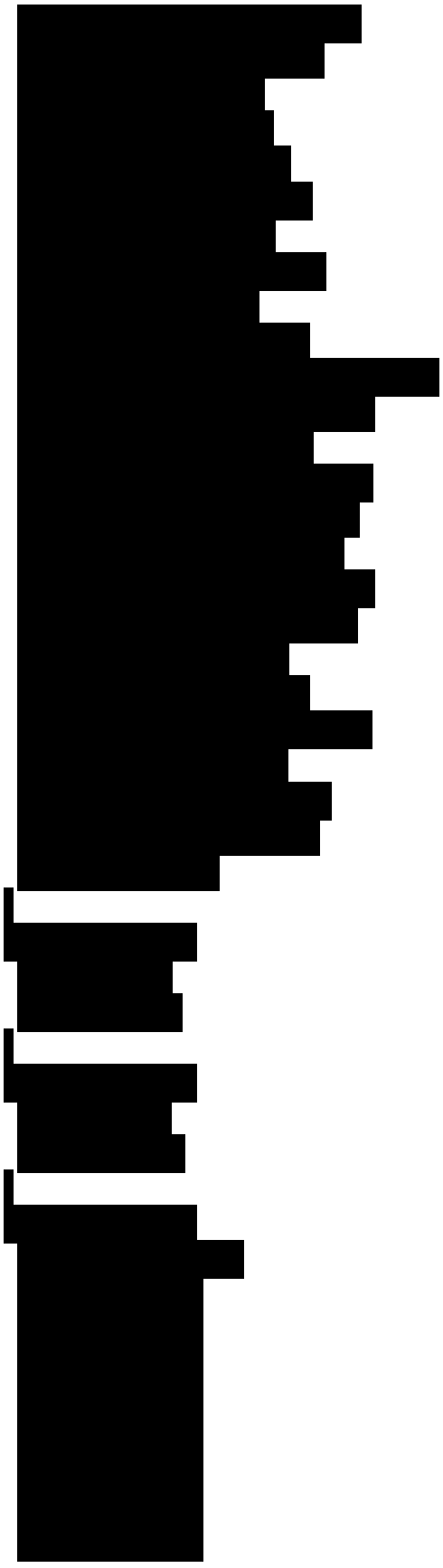
GOOG-BRWN-00840745 Row 208

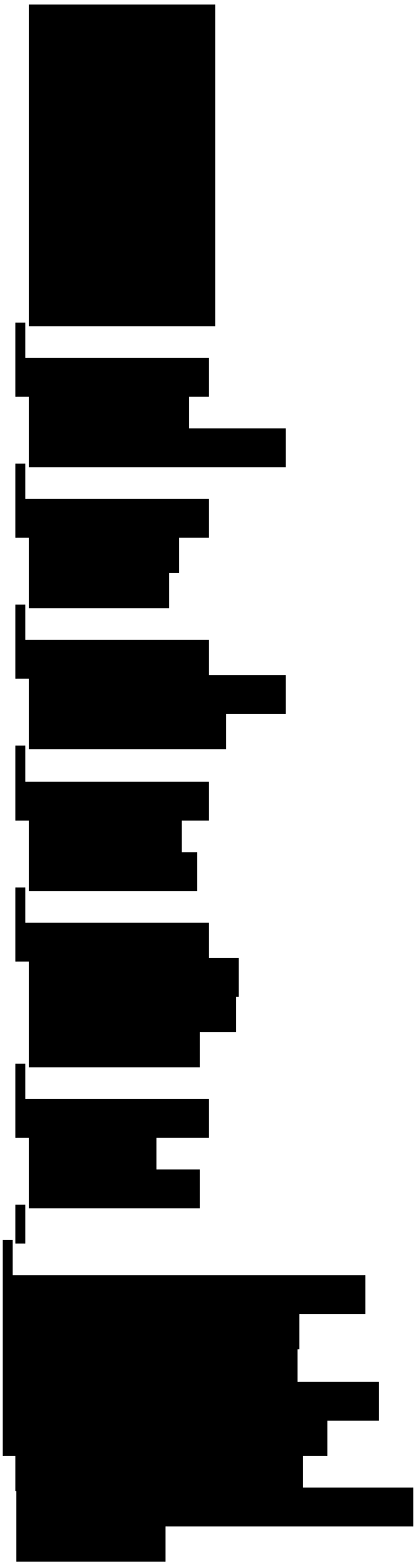


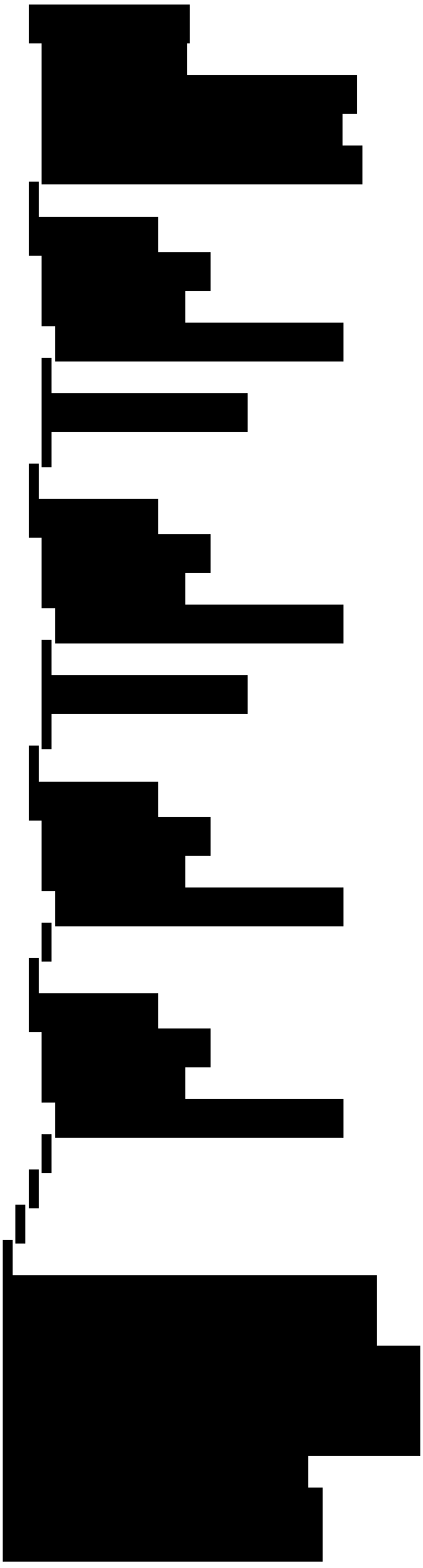


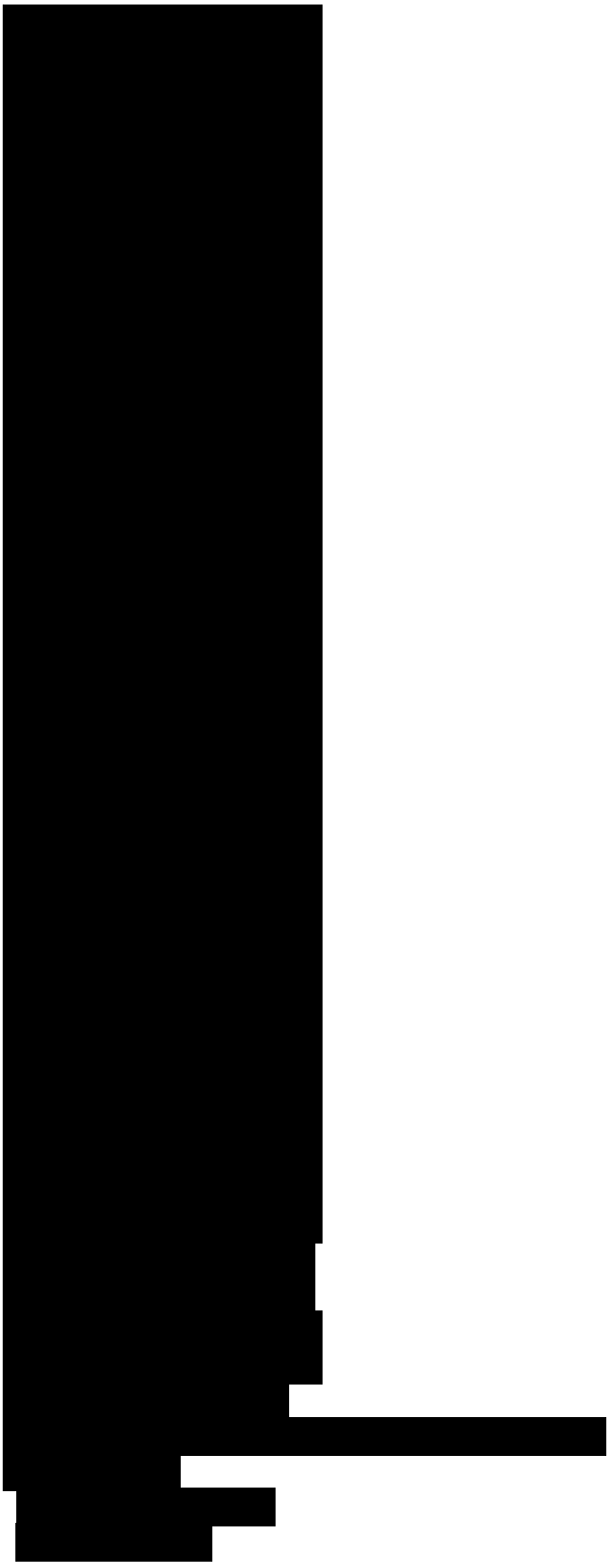


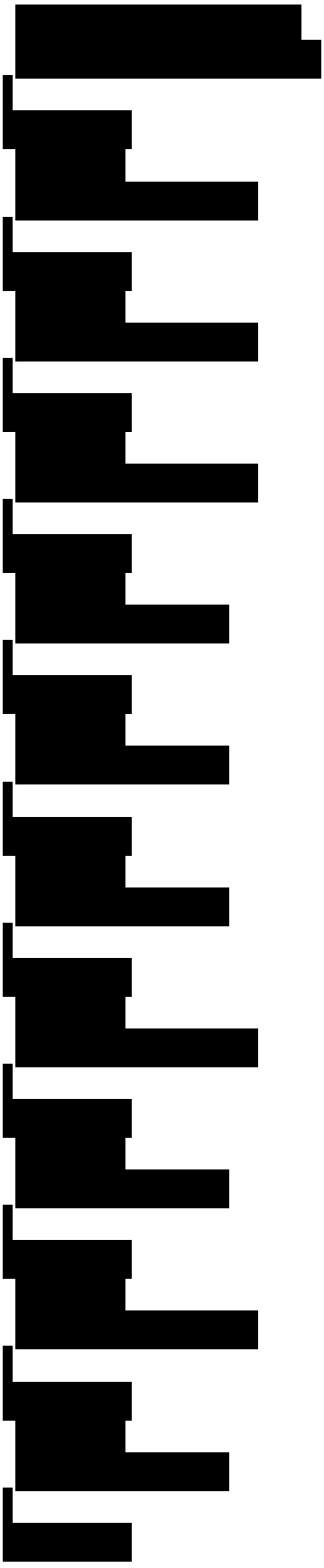


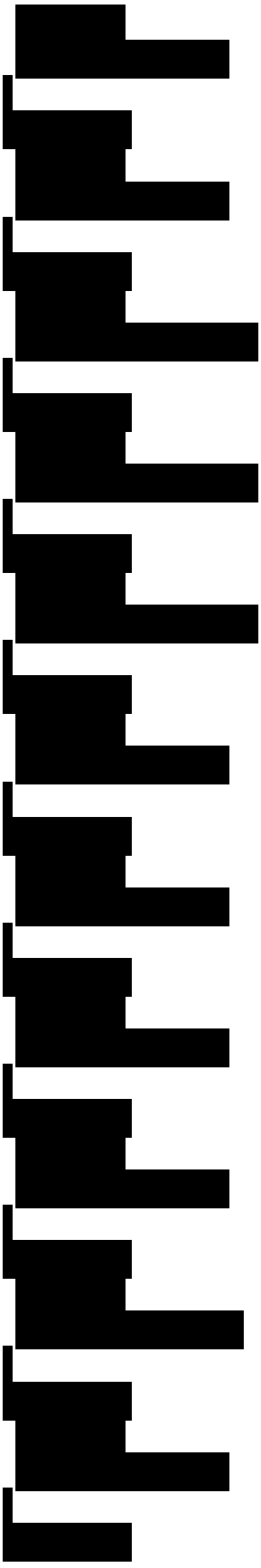


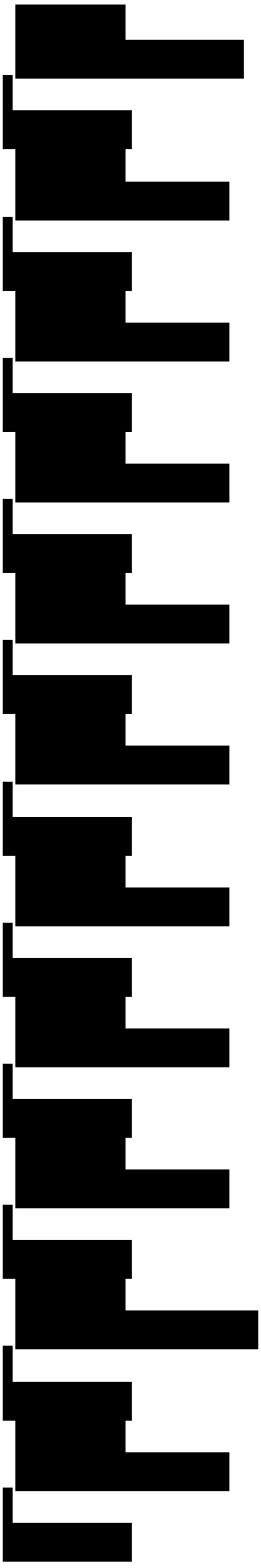


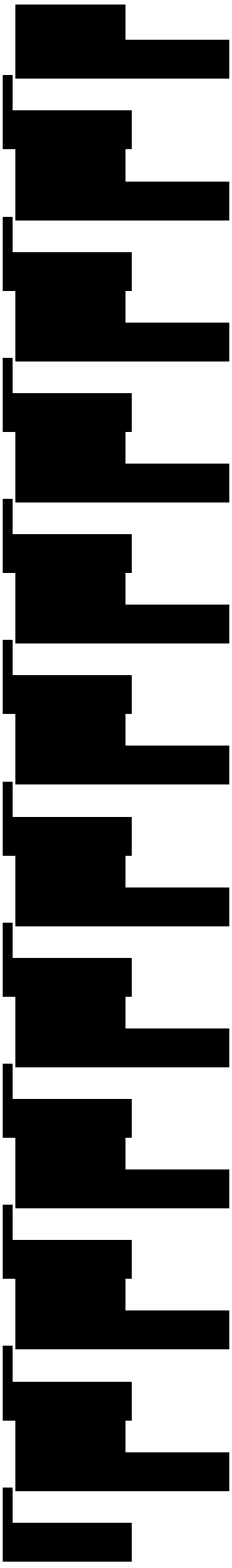


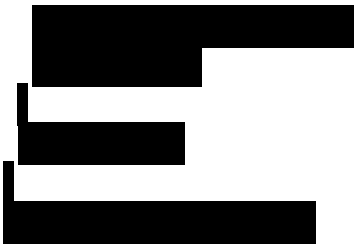












EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

EXHIBIT D

Produced Electronically

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

EXHIBIT E

Produced Electronically

EXPERT REPORT OF JONATHAN E. HOCHMAN

APRIL 15, 2022

EXHIBIT F

Produced Electronically